图像视频语义分析

薛向阳

xyxue@fudan.edu.cn

计算机科学技术学院

复旦大学

致谢：研究团队主要成员

姜育刚，副教授	路红，副教授	张巍，副教授	张珂，研究生	陆遥，研究生
				
视频动作识别	视频结构分析	图像标注	图像语义分割	图像显著性检测

目录
Contents

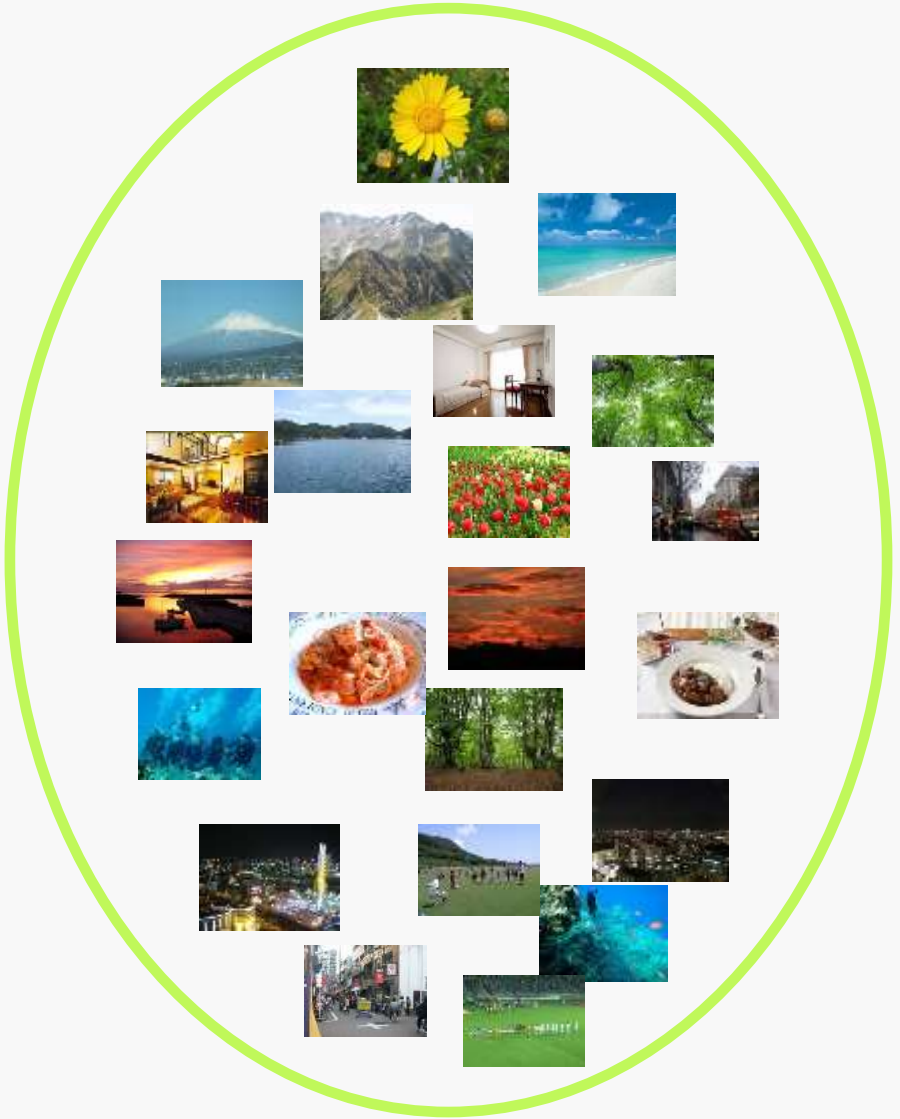
一

研究背景

- 互联网上有数千亿幅照片
- 全世界有数亿摄像头

互联网用户上传了数千亿幅照片！

随时随地拍照，方便快捷发布！



如何实现互联网上海量图片的共享？

--- 管理、搜索、推荐

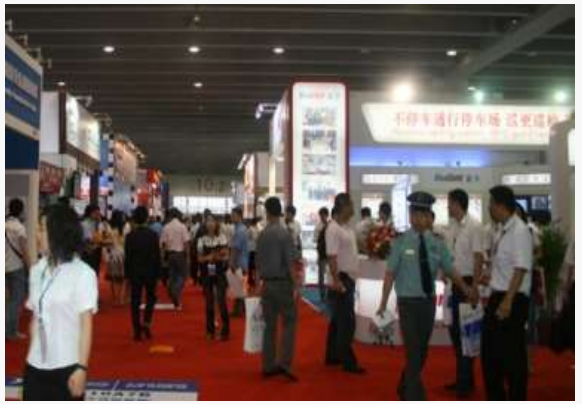
6

全世界安装了数亿摄像头!

道路/交通要道监控

治安/破案/取证

码头/机场/车站监控



展馆/会场监控

企业运营监控

金融银行监控

小区/家庭监控

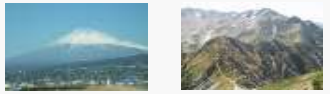
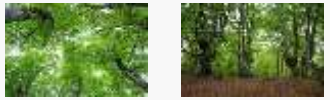
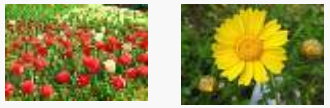
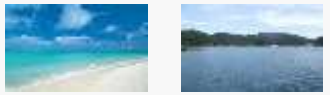
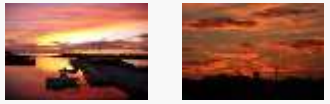
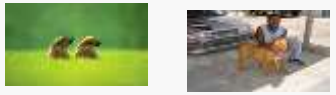
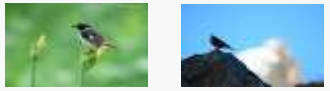
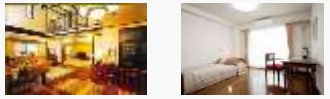
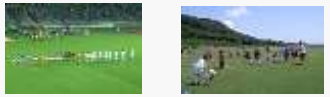
二

什么是语义分析

- 给**图像**一个类别或多个关键词
- 给**图像**多个关键词，并指出其所在区域
- 给**图像**一句或多句话描述
- 给**视频**一个动作类别或语句描述

图像分类：给图像一个类别标签

分类

类别	图像
Mountain	
forest	
flower	
sea	
sunset	
dog	
bird	
indoor	
sport field	

+ 目标类（Object）

Sheep

Boat

+ 场景类（Scene）

Indoor

Party

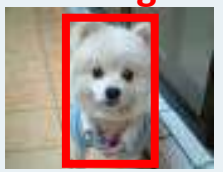
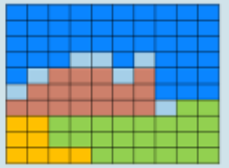
Wedding

Street

图像标注：给图像多个关键词

标注

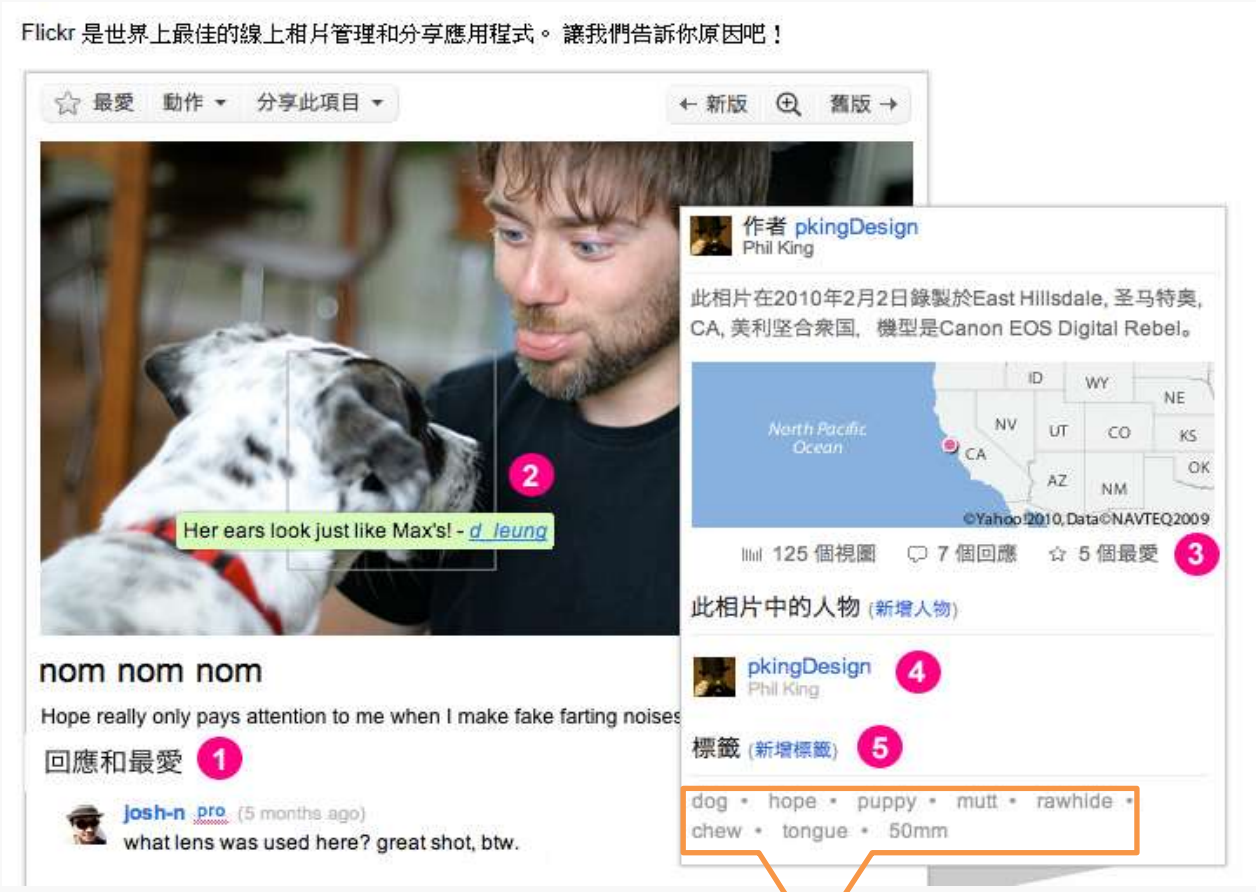
图像	关键词
	mountain sky
	Flower
	Sea, sky, Beach
	sunset sea
	street building road
	indoor chair window

标注方式	图像样例	标注强度
整幅图像	Melon  Beach 	弱
外接矩形	Dog  Cat 	
均匀网格	  Sky Rock Water Sand	
像素精度	  Dog Ground	强

图像标注：共享网站（如Flickr、Facebook等）

Social Images with Tags

- Tags are words or phrases
- Tags are searchable within the site, and can show popular topics
- Tags improve search relevance

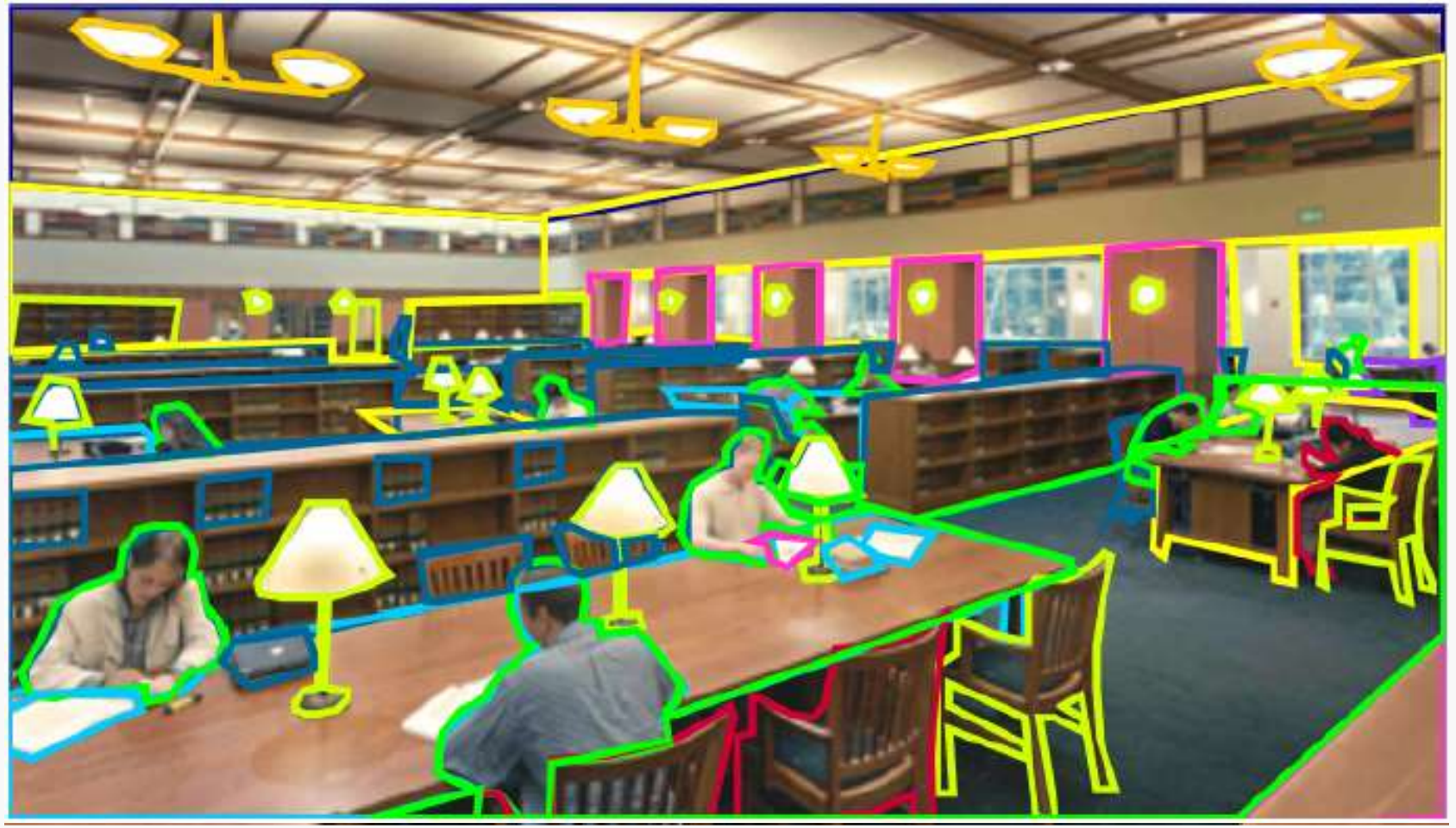


TAGS: dog, hope, puppy, mutt, rawhide, chew, tongue, 50mm

Tags are always noisy, subjective, synonymous, …

区域标注：给图像多个关键词，并指出关键词所在区域

人工标注也是一件很困难的事情!



Notes on image annotation, Adela Barriuso, Antonio Torralba, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology

看到这幅照片，您还想标注吗?

Notes on image annotation, Adela Barriuso, Antonio Torralba, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology

图像标注：给图像一句或几句话

图书馆灯火通明，有10多人在图书馆里学习...

视频标注：给视频一个动作类别或几句话

Hollywood2
- 1707 clips
- 12 classes


GetOutCar


Kiss


Run


FightPerson


Stand Up

Olympic Sports
-783 clips
-16 classes


Gymnastics Vault


Diving Platfrom


High Jump


Javelin Throw


Long Jump

HMDB51
-6766 clips
- 51 classes


Dive


Dribble


Hug


Ride Bike


Shake Hands

三

代表性方法介绍

- 原理框图
- 图像表示
- 目标检测

图像标注：原理框图

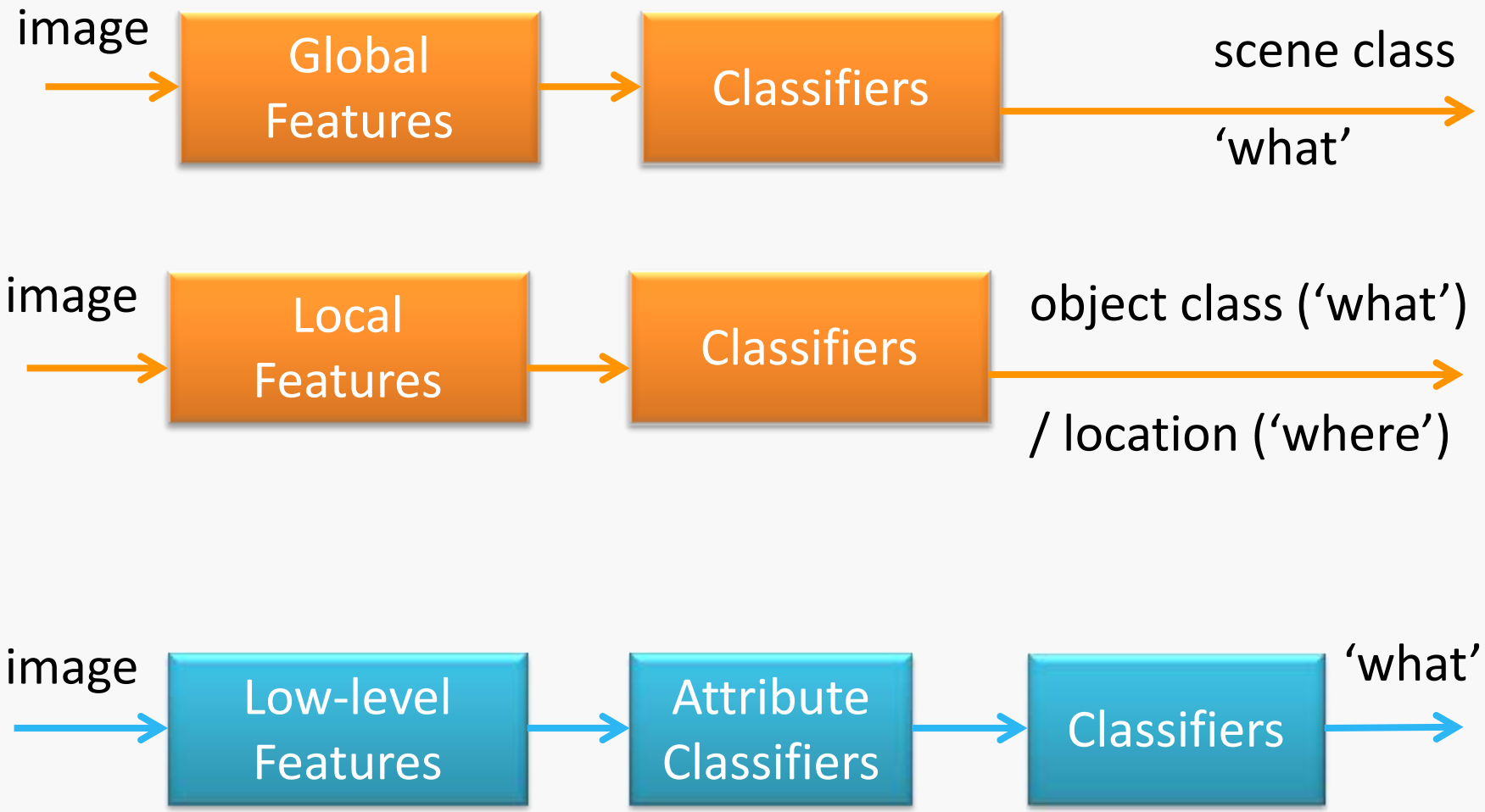
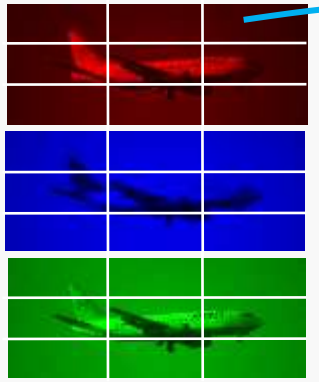


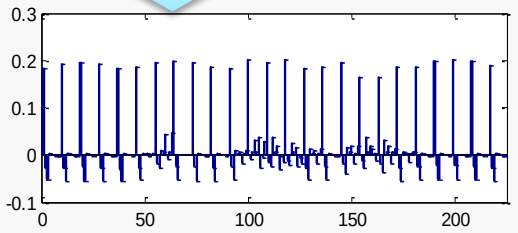
image
Global Features
Classifiers
scene class
'what'
image
Local Features
Classifiers
object class ('what')
/ location ('where')
image
Low-level Features
Attribute Classifiers
Classifiers
'what'

图像表示：全局特征描述图像内容

Grid Color Moment

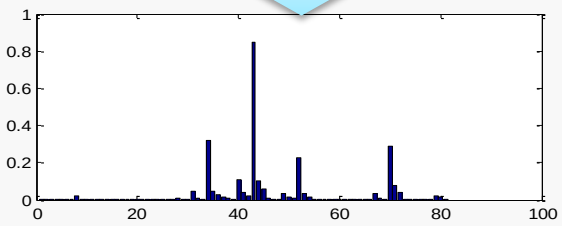


mean
variance
…

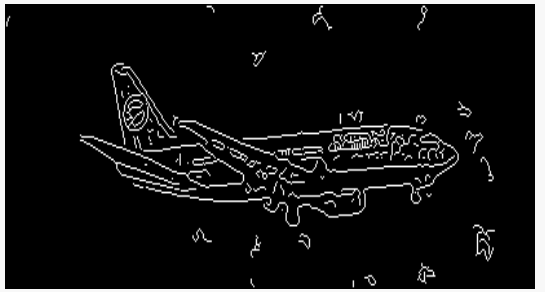


Wavelet Texture

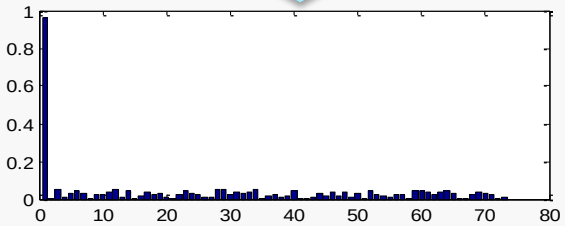
energy in different filter banks



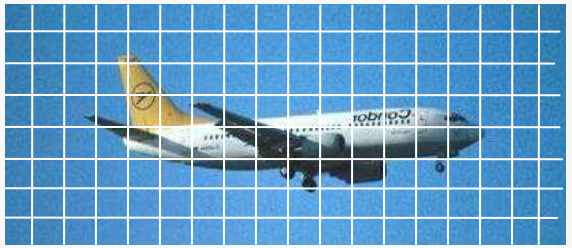
Canny Edge

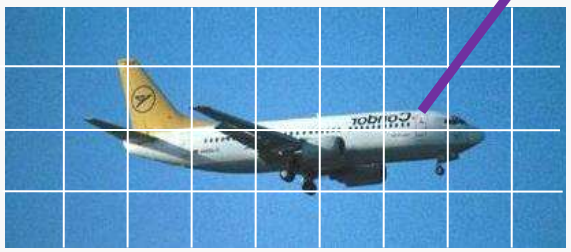


edge direction counts

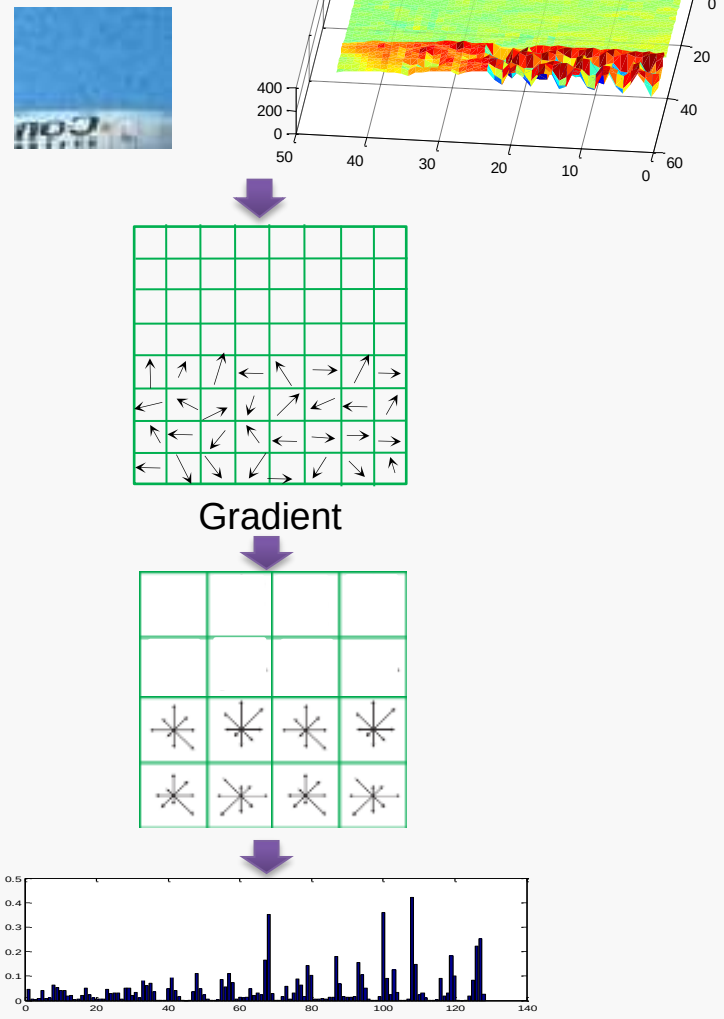


图像表示：局部特征描述图像内容

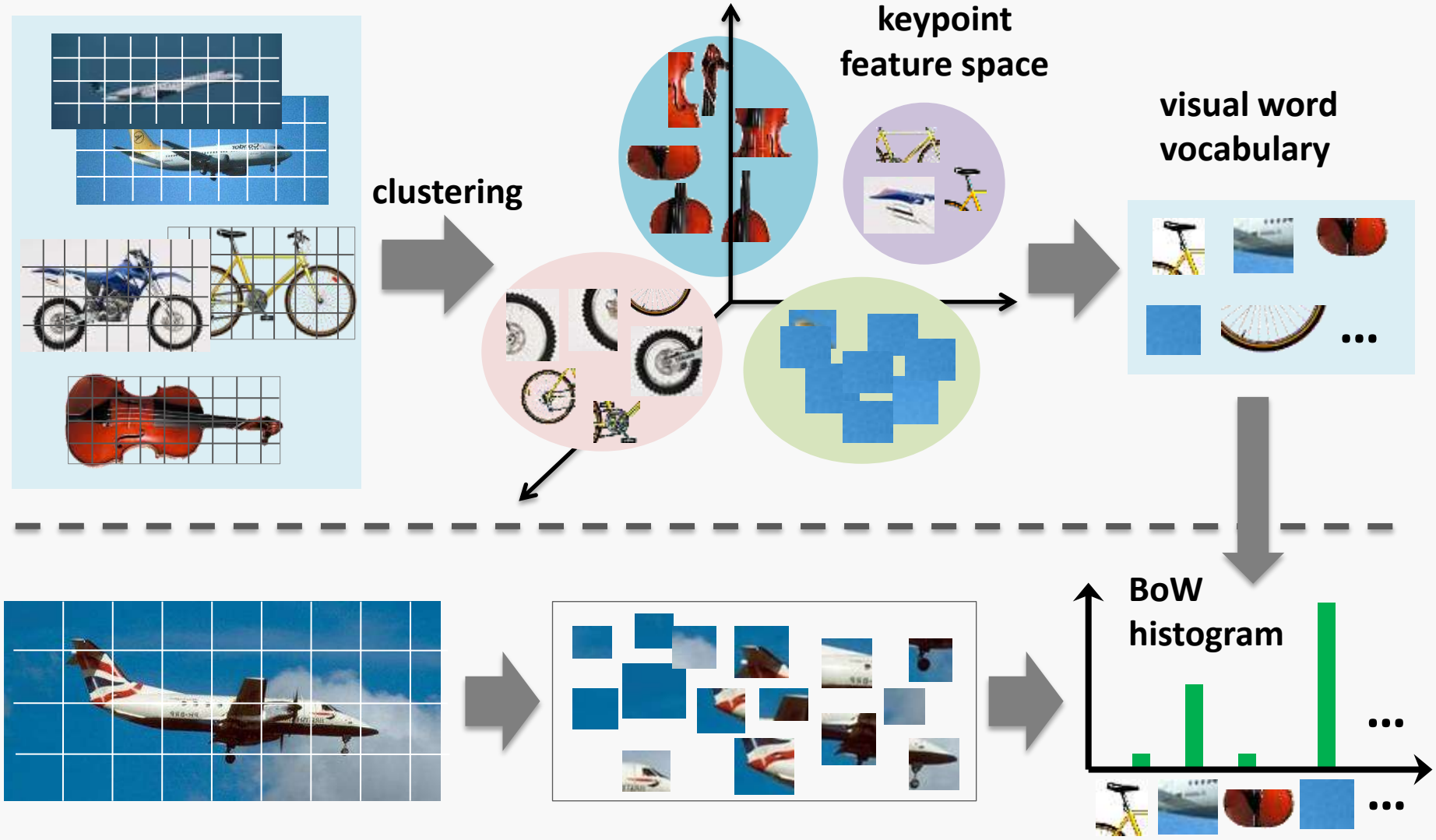




D. G. Lowe, IJCV 2004

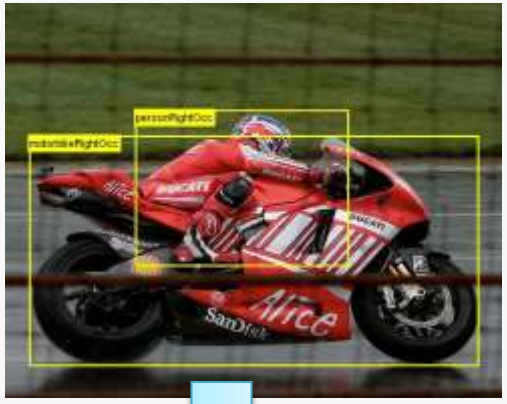


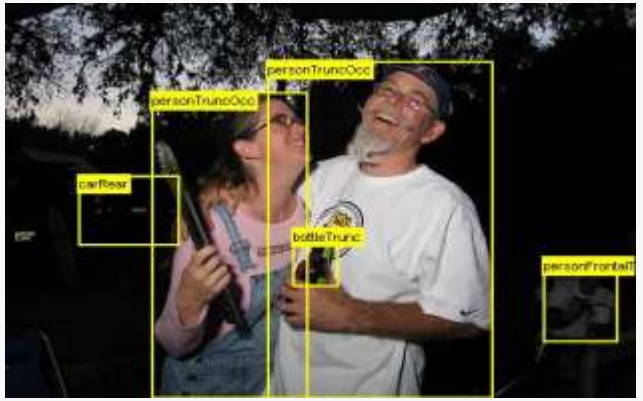
图像表示：视觉词袋描述图像内容



Shih-Fu Chang, Keynote Talk@ ICME09

目标检测：搜索标签所在区域





Training images are labeled with object bounding box, marked with truncated/occlusion/difficult

Multiple classifiers for objects with different viewpoints/aspect ratio

- Sliding window with multiple size/ratio according to training images
- Similar features (global/local) extracted from sub-image for classification

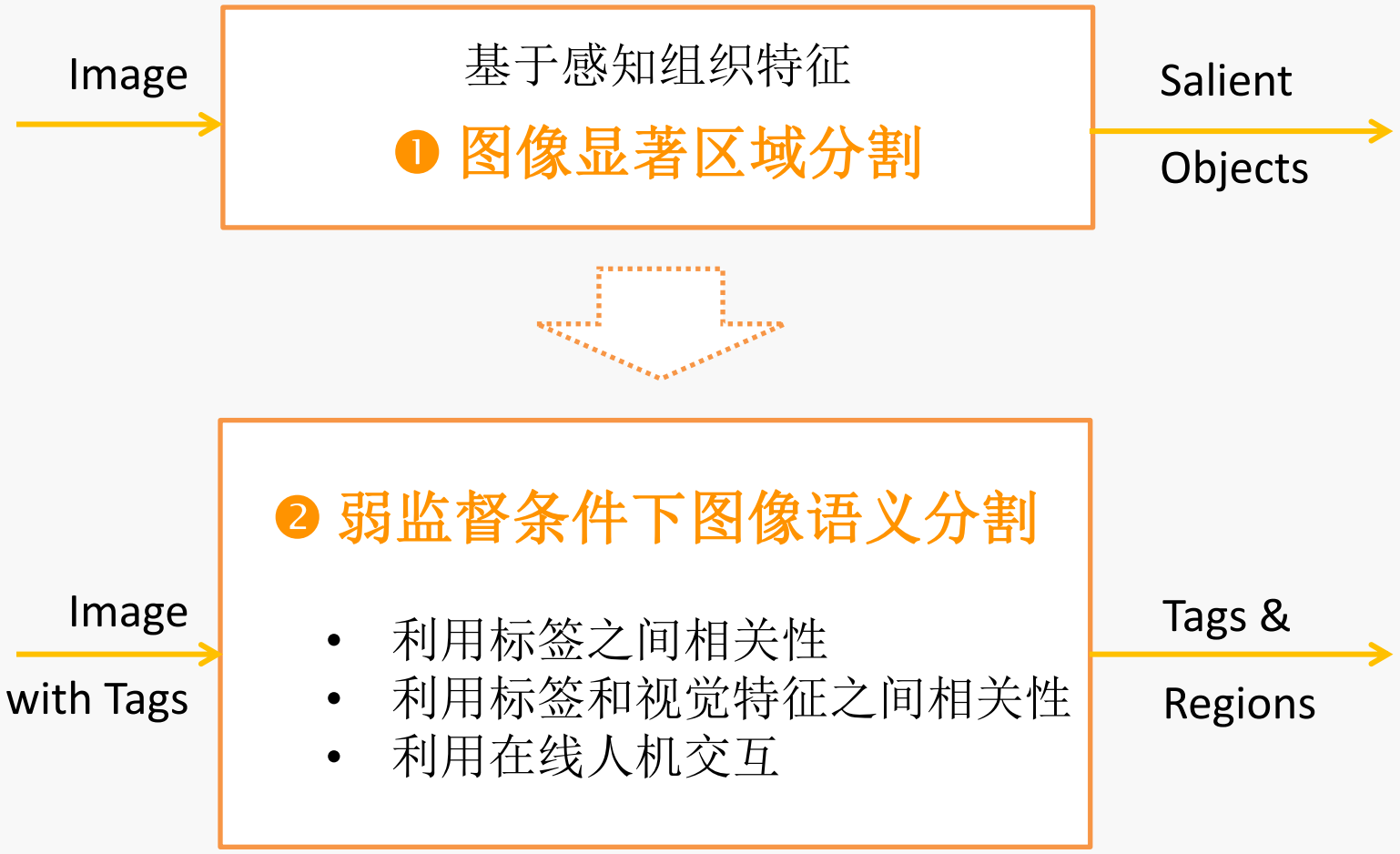
Vedaldi, Varma, Gulshan & Zisserman, ICCV 2009

四

我们的研究进展

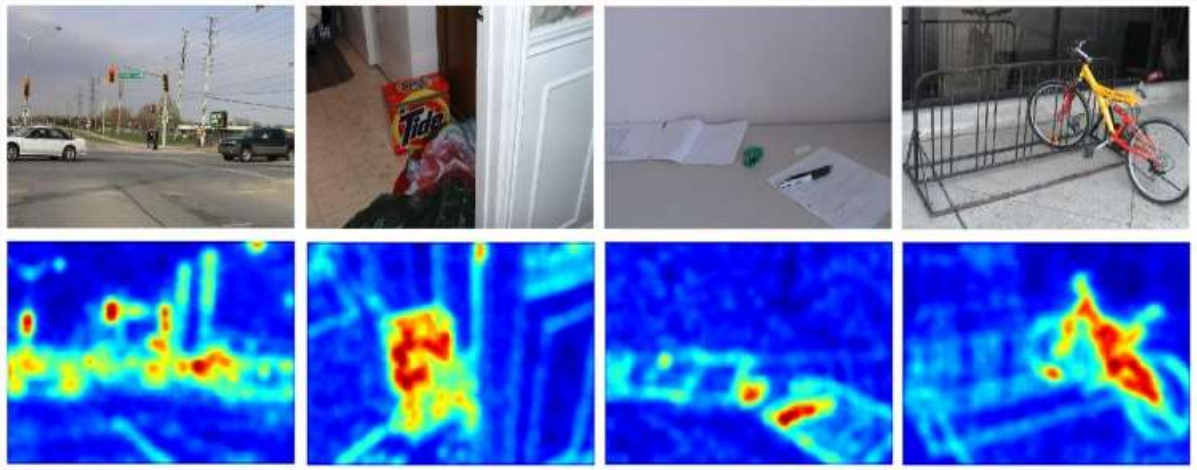
- 显著区域分割
- 图像语义分割
- 视频动作识别

图像分割与标注：研究思路



显著区域检测：在图像中寻找引人注意的区域

基于像素的显著性检测方法



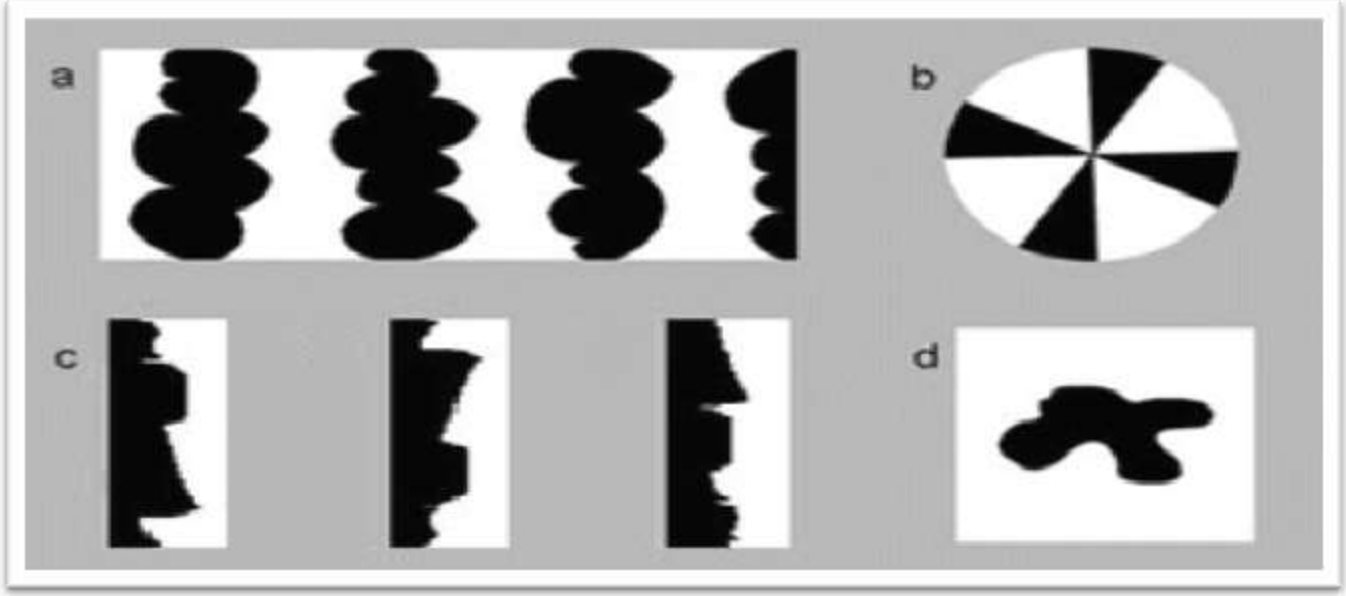
Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.

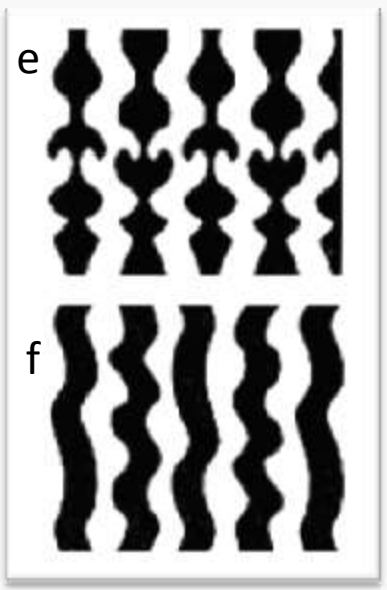
基于区域的显著性检测方法

Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection. 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011: 409-416.

显著区域检测：感知组织特征简介

感知组织（Perceptual Organization）包括感知分割（Perceptual Segmentation）及前景-背景组织（Figure-Ground Organization）。



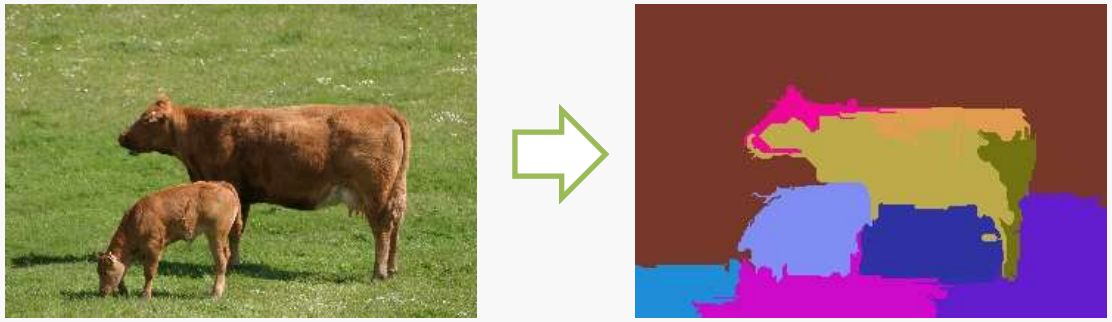


感知组织特征：（a）凸性，（b）尺寸大小，（c-左）熟悉物体，（d）包围性，（e）对称性和（f）平行性等等。**感知组织特征的计算**是计算机视觉的重要研究课题，它对显著性区域检测、图像分割、物体识别和图像检索等任务都具有重要价值。

图片来源：http://www.scholarpedia.org/article/Figure-ground_perception

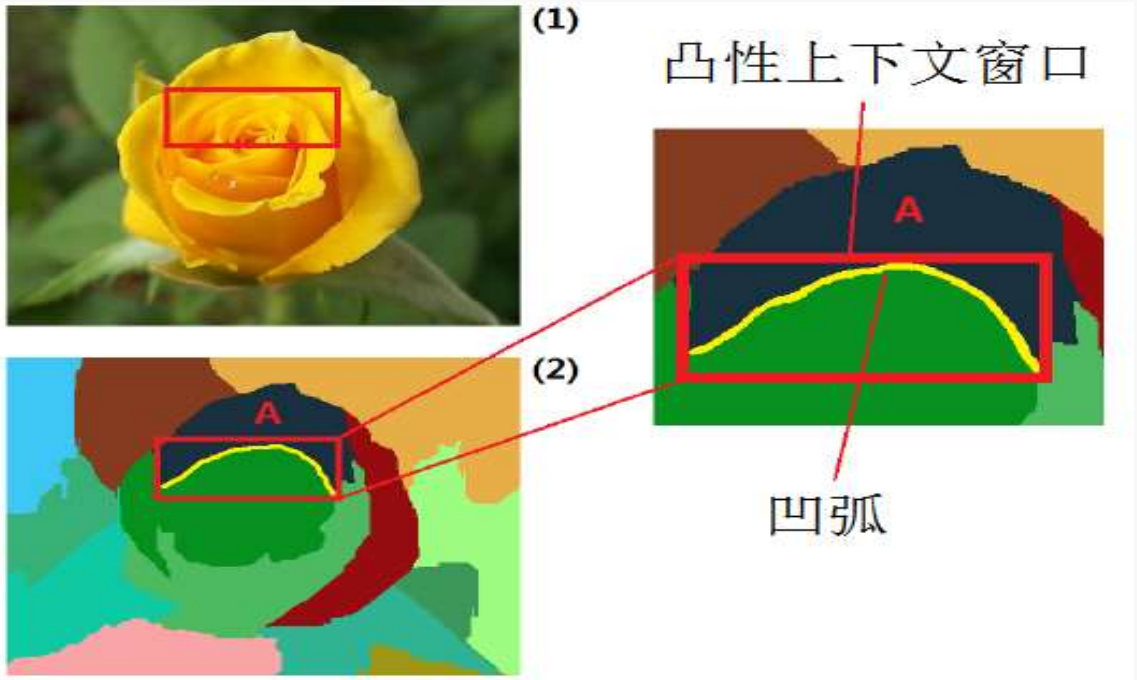
感知组织特征计算：用区域边界计算凸性(convexity)

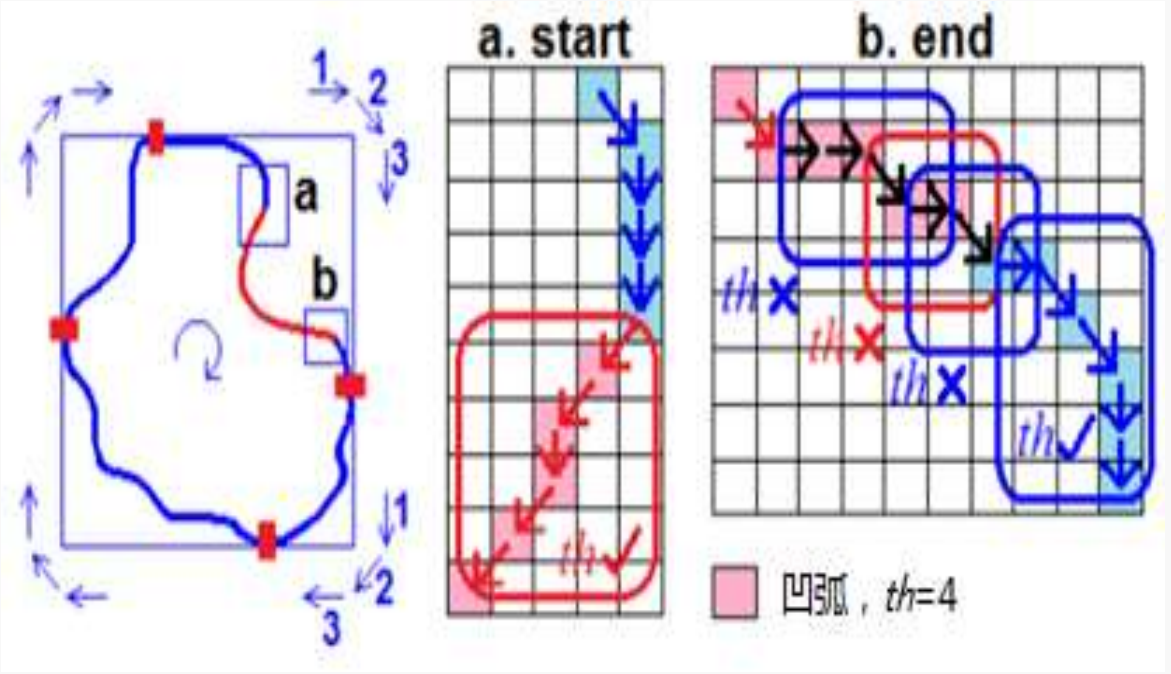
区域分割



Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation. International Journal of Computer Vision, 2004, 59(2): 167-181.

检测矩形窗口内凹弧

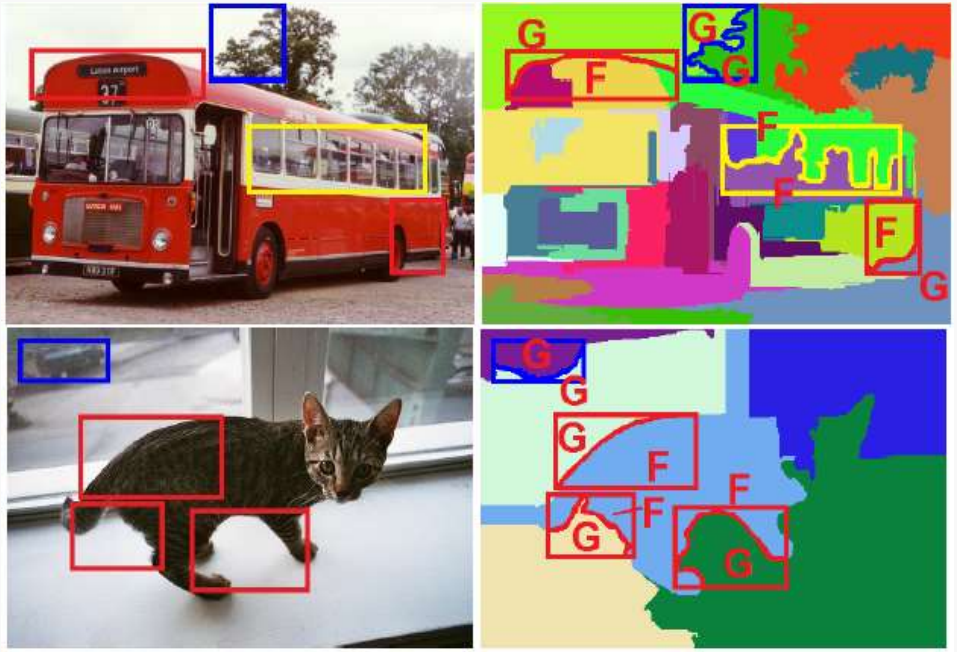




凸性检测算法

Yao Lu, Wei Zhang, Hong. Lu, Xiangyang Xue. Salient Object Detection using Concavity Context. 13th IEEE International Conference on Computer Vision. Barcelona, Spain. 2011.

感知组织特征计算：用自然图像检验凸性的有效性



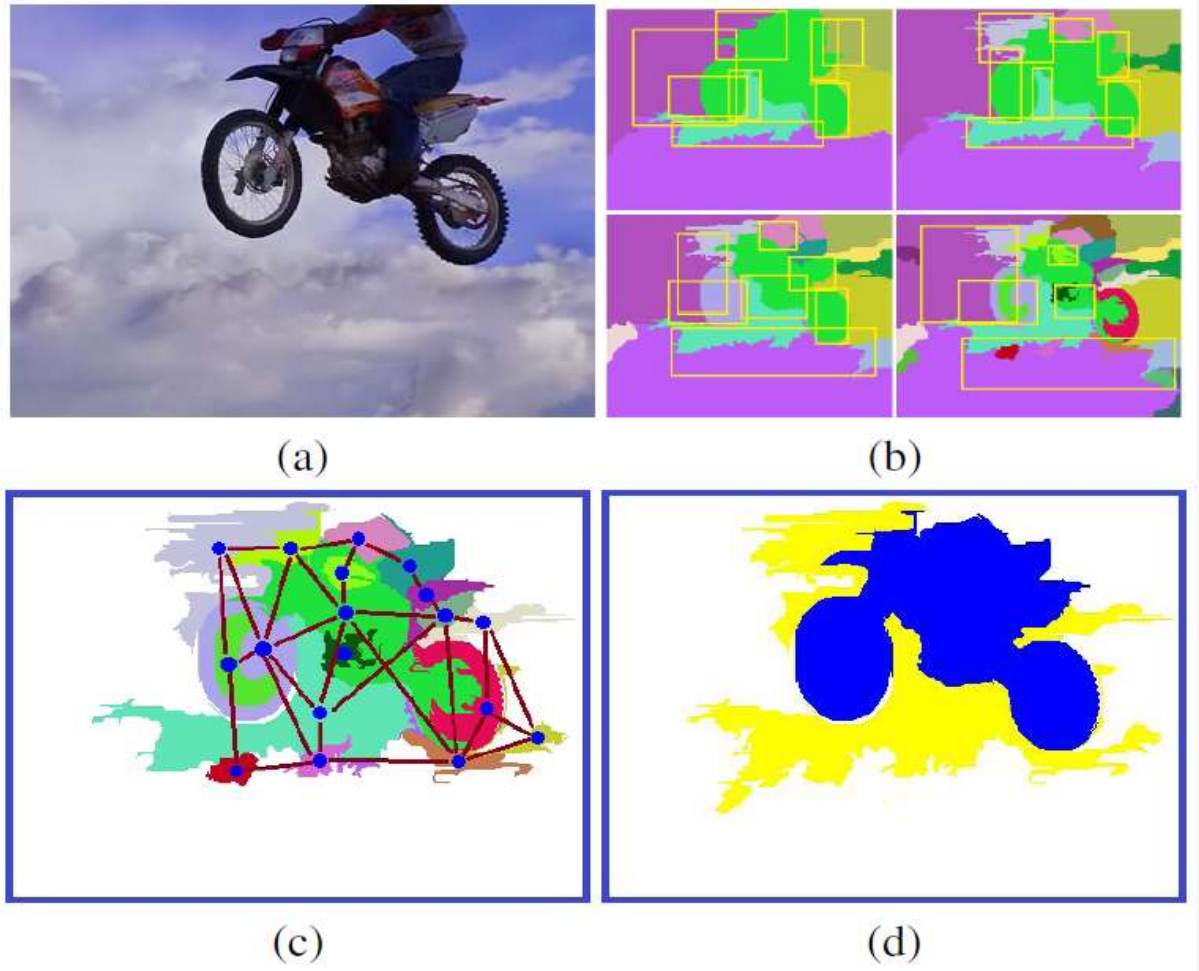
Dataset	MSRC	VOC08	SUN	Avg
Case 1. F-F	48/19%	114/21%	112/26%	22%
Case 2. G-G	46/19%	124/23%	69/16%	19%
Case 3. F-G	154/62%	297/56%	254/58%	59%
Total #	248	535	435	/
Avg CCW Size	29%×27%	21%×22%	23%×21%	/
Chance	66%	58%	57%	60%

F -F：窗口内凹弧两边都是前景
G-G：窗口内凹弧两边都是背景
F -G：窗口内凹弧两边分别是背景与前景

在三个数据集240幅自然图像上进行检验，81%的窗口含有前景物体（F-F + F-G）。在随机选取窗口中，60%含有前景物体。可见，“凸性”对前景检测有一定分辨能力，但不是压倒性的！

Yao Lu, Wei Zhang, Hong. Lu, Xiangyang Xue. Salient Object Detection using Concavity Context. 13th IEEE International Conference on Computer Vision. Barcelona, Spain. 2011.

显著区域分割：利用凸性特征改变图中边的权重



分割算法：

a. 首先对图像进行多分割，即调整分割尺度（最小面积），同时保持其它参数不变，可在不同尺度上分割图像，以提高最终分割精度；

b. 在所有尺度上检测凹弧及凸性上下文窗口，将所有凸性上下文窗口映射到分割粒度最小的尺度上；

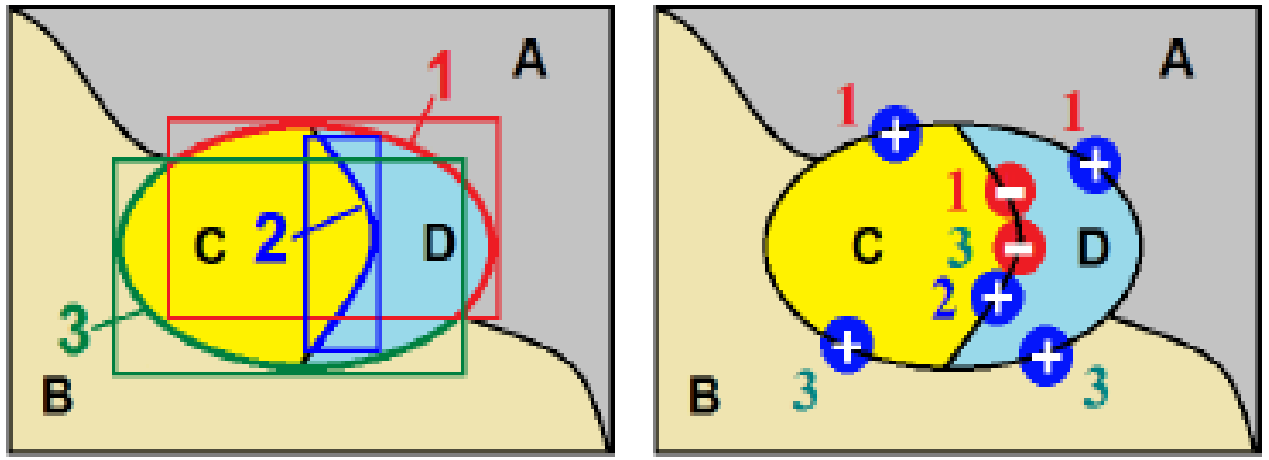
c. 构建一个图，其中所有凸性上下文窗口所覆盖的超像素对应到图的一个结点，结点之间距离（权重）用凸性特征及层次树信息进行计算；

d. 对图进行二分割（NCut），可得到显著区域或前景-背景的分割结果！

Yao Lu, Wei Zhang, Hong. Lu, Xiangyang Xue. Salient Object Detection using Concavity Context. 13th IEEE International Conference on Computer Vision. Barcelona, Spain. 2011.

如何利用凸性特征改变图中边的权重?

权重调整示意图

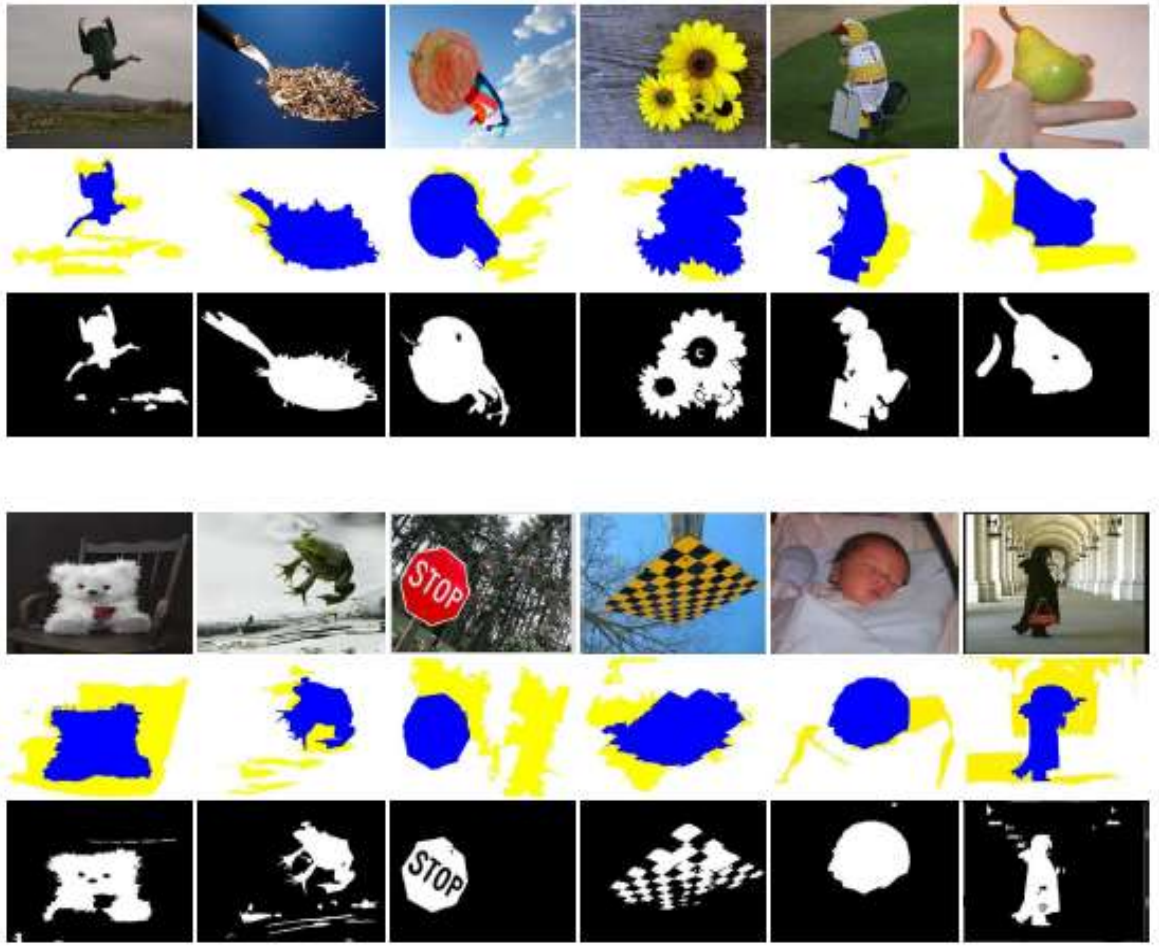


如左图所示，假设CD是前景，AB是背景，ABCD之间原始距离都为1。我们的目的是将AB之间距离减小，CD之间的距离减小，并使AB、CD之间的距离增加（A与C，A与D，B与C，B与D）。

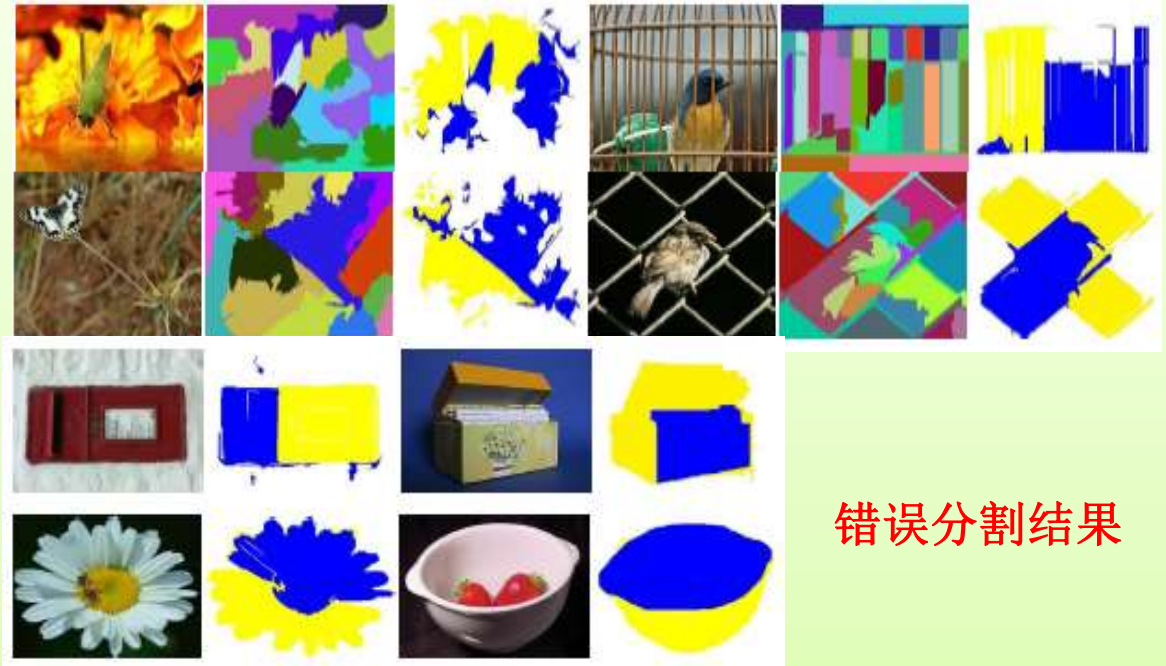
- 对凹弧1，考察其两侧的图像区域集合，$S_1^+ = \{A\}$，$S_1^- = \{C, D\}$。于是，W(C,D)=>c=c-1；W(A,C)=>c=c+1；W(A,D)=>c=c+1。
- 对凹弧2，考察其两侧的图像区域集合，$S_2^+ = \{C\}$，$S_2^- = \{D\}$。于是，W(C,D)=>c=c+1。
- 对凹弧3，考察其两侧的图像区域集合，$S_3^+ = \{B\}$，$S_3^- = \{C, D\}$。于是，W(C,D)=>c=c-1；W(B,C)=>c=c+1；W(B,D)=>c=c+1。
- 调整后：*W*(C,D)中c= -1，即CD之间距离被减小；而*W*(A,D) *W*(A,C) *W*(B,D) *W*(B,C)中c=1，即它们之间的距离被增加了，达到了距离调整的目的。

Yao Lu, Wei Zhang, Hong. Lu, Xiangyang Xue. Salient Object Detection using Concavity Context. 13th IEEE International Conference on Computer Vision. Barcelona, Spain. 2011.

实验结果：利用凸性特征提升分割性能



蓝色为前景物体，黄色及白色为背景



错误分割结果

背景信息过强，导致无法对前景物体正确分割；前景物体本身具有很强层次性，并含有较少背景，导致分割时将前景分为多个部分

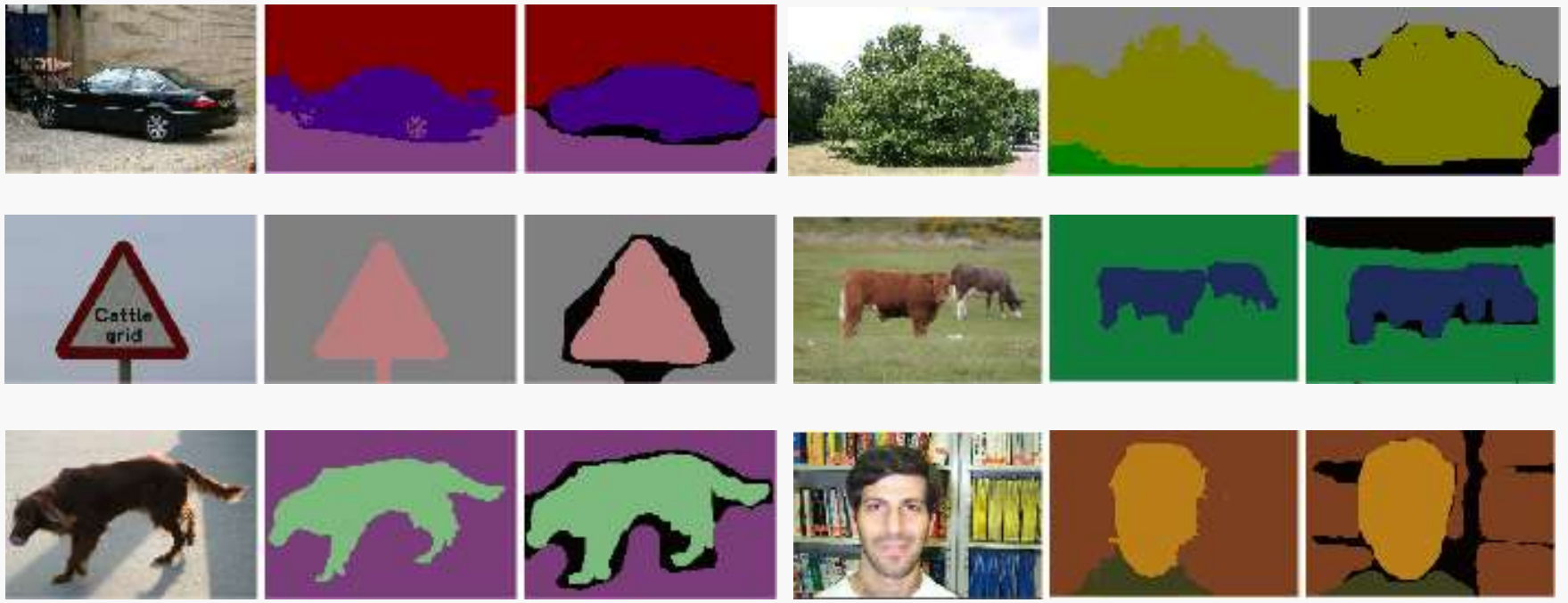
Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection. IEEE Conference on Computer Vision and Pattern Recognition, 2009: 1597-1604.

Yao Lu, Wei Zhang, Hong. Lu, Xiangyang Xue. Salient Object Detection using Concavity Context. 13th IEEE International Conference on Computer Vision. Barcelona, Spain. 2011.

显著区域分割：进一步工作是利用更多感知组织特征

- 我们仅用了凸性感知组织特征和无监督（聚类）方法，就能较好提升显著区域（前景）检测（分割）精度；
- 在CVPR 2012论文 ***Learning Attention Map from Images*** 中，我们采用更多的感知组织特征（如凸性、包围性、对称性等）和有监督学习方法，进一步提升了检测性能；
- 显著区域分割方法并不能指出该显著区域是什么语义概念（what）。

图像语义分割：将图像分割为主要区域，并赋于标签



Ke Zhang, Wei Zhang, Yingbin Zheng, Xiangyang Xue, Sparse Reconstruction for Weakly Supervised Semantic Segmentation, IJCAI 2013

图像语义分割：传统方法及其不足；研究动机

- 需要大量像素级人工标注训练数据！费时费力，训练数据很难上规模！

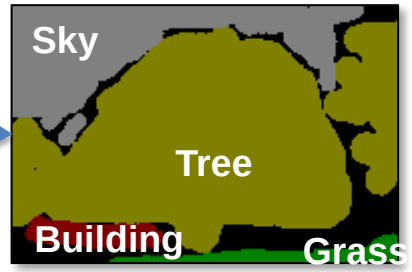
训练图像集

Training

Classifier

测试图像

Testing



Results

- Semantic Texton Forest [Shotton et al., CVPR2008]

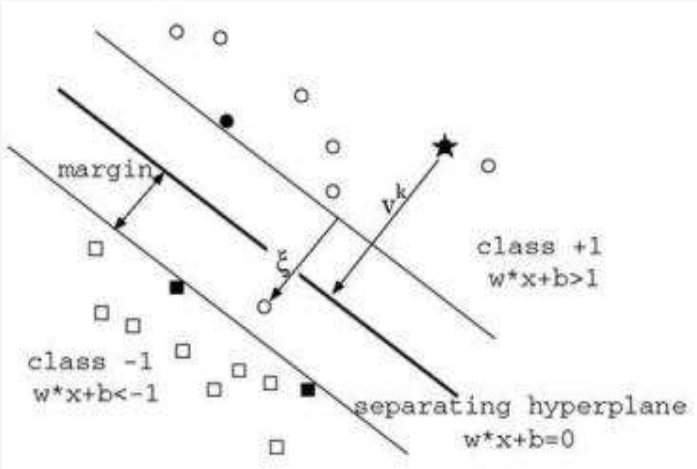
事实上，我们已经有了海量的附带语义标签的图像数据集

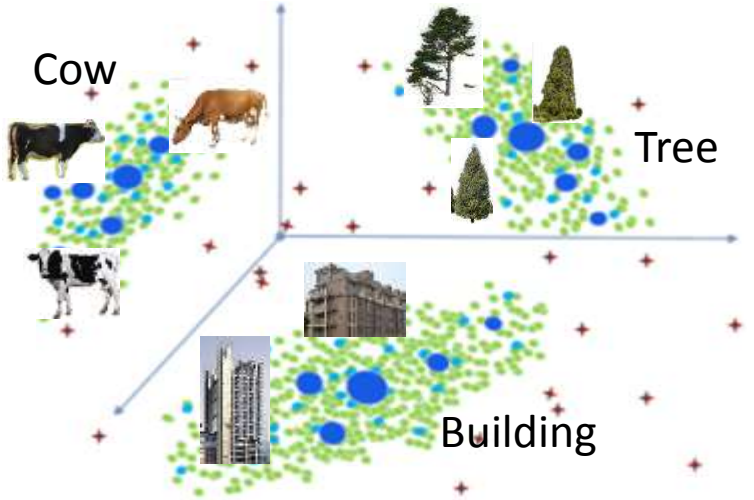
Building
Sky
Grass
Tree

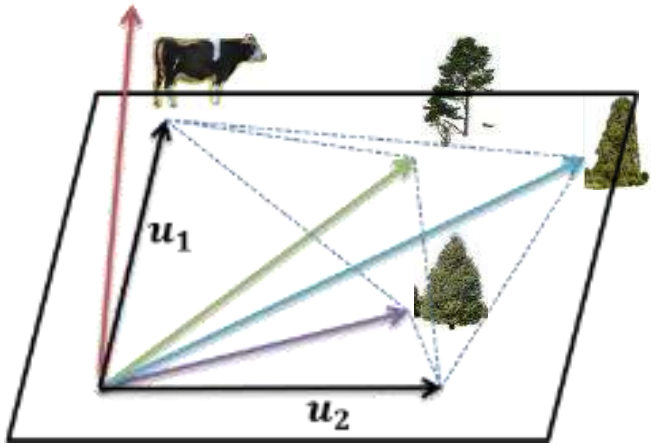
有没有办法直接利用这些数据？（可大大节省人工标注成本，极易获得海量的训练图像数据集）

Ke Zhang, Wei Zhang, Yingbin Zheng, Xiangyang Xue, Sparse Reconstruction for Weakly Supervised Semantic Segmentation, IJCAI 2013

图像语义分割：准备知识与假设







SVM分类器由模型参数（w，b）确定。假如能定义一个合理的评价分类器好坏的代价函数score(w,b)，那么就可以估计出某种意义下的最佳参数（w*，b*）

在高维特征空间中，同一类别的图像区域的视觉特征向量应嵌入在同一个低维子空间中

$$\min_W f(W) = (\|X - XW\|_F)^2 + \alpha(\|W\|_1)^2$$
$$s.t.\ \mathbf{1}^\top W = \mathbf{1}^\top,\ W_{ij} \in [0,1],\ (i,j = 1,\ldots,N)$$

同一类别的图像区域的视觉特征向量，用该类别所在子空间的基矢量重构时，误差很小，否则，误差很大（期望特征具有很好的描述及区分能力）

Ke Zhang, Wei Zhang, Yingbin Zheng, Xiangyang Xue, Sparse Reconstruction for Weakly Supervised Semantic Segmentation, IJCAI 2013

图像语义分割：为某一个语义标签训练分类器

例：为Tiger训练分类器

图像数据集

无tiger的图像集

有tiger的图像集

获取负样本特征

超像素区域分割 → 特征提取

获取正&负样本特征

超像素区域分割 → 特征提取

SVM（w’，b’）分类器

用分类器预测的正样本稀疏构建子空间（基）

计算负样本的重构误差

随机参数采样

分类器评估

参数优化

重构误差越大，则表明分类器越好！？

图像语义分割：我们的方法

训练图像集
附标签

超像素

在分类器模型参数空间中采样

$$(w_1, b_1), (w_2, b_2), \ldots, , (w_m, b_m)$$

采用子空间重构法评估出最优分类器（SVM）

- 充分利用互联网上海量带标签图像，标签即意味图像中主要有什么语义
- 用子空间重构法评估分类器好坏

Training

Testing

测试图像

每类一个最佳分类器

语义分割结果

图像语义分割：实验结果

MSRC Dataset 像素级分割精度（%）

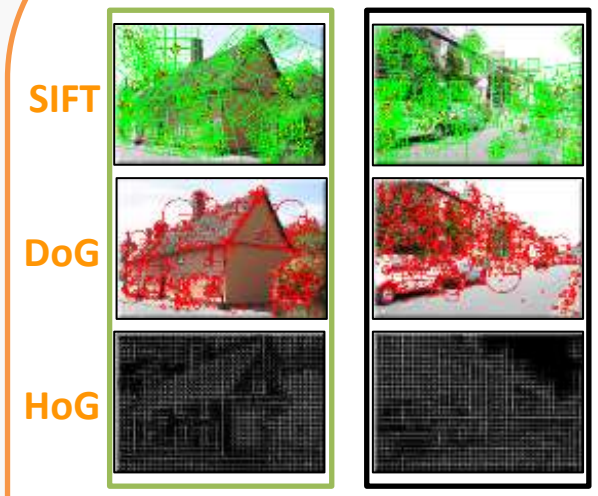
Methods	building	grass	tree	cow	sheep	sky	aeroplane	water	face	car	bicycle	flower	sign	bird	book	chair	road	cat	dog	body	boat	average
[Shotton 2006]	62	**98**	86	58	50	83	60	53	74	63	75	63	35	19	92	15	86	54	19	62	07	58
[Yang 2007]	63	**98**	90	66	54	86	63	71	83	71	80	71	38	23	88	23	88	33	34	43	**32**	62
[Shotton 2008]	49	88	79	**97**	**97**	78	82	54	**87**	74	72	74	36	24	93	51	78	75	35	**66**	18	67
[Ladicky 2009]	**80**	96	86	74	87	**99**	74	**87**	86	87	82	**97**	**95**	30	86	31	**95**	51	**69**	**66**	09	75
[Csurka2011]	75	93	78	70	79	88	66	63	75	76	81	74	44	25	75	24	79	54	55	43	18	64
[Lucchi 2012]	59	90	**92**	82	83	94	**91**	80	85	**88**	**96**	89	73	**48**	**96**	**62**	81	**87**	33	44	30	**76**
[Verbeek2007]	45	64	71	75	74	86	81	47	1	73	55	**88**	6	6	63	18	80	27	26	55	8	50
[Vezhnevets 2010]	7	**96**	18	32	6	**99**	0	46	**97**	54	74	54	14	9	82	1	28	47	5	0	0	37
[Vezhnevets 2011]	5	80	58	**81**	**97**	87	**99**	63	91	**86**	**98**	82	67	**46**	59	45	66	**64**	**45**	33	**54**	67
Ours	**63**	93	**92**	62	75	78	79	**64**	95	79	93	62	**76**	32	**95**	**48**	**83**	63	38	**68**	15	**69**

bold = winner, the upper methods are fully supervised and the lowers are weakly supervised

总结：巧妙利用带标签图像数据训练语义计算模型，实现了自动语义分割，不仅可以获得图像的主要区域，还能给区域赋予语义标签！

图像标注：标签补全

多种高维特征



语义标签补齐

building
sky
tree

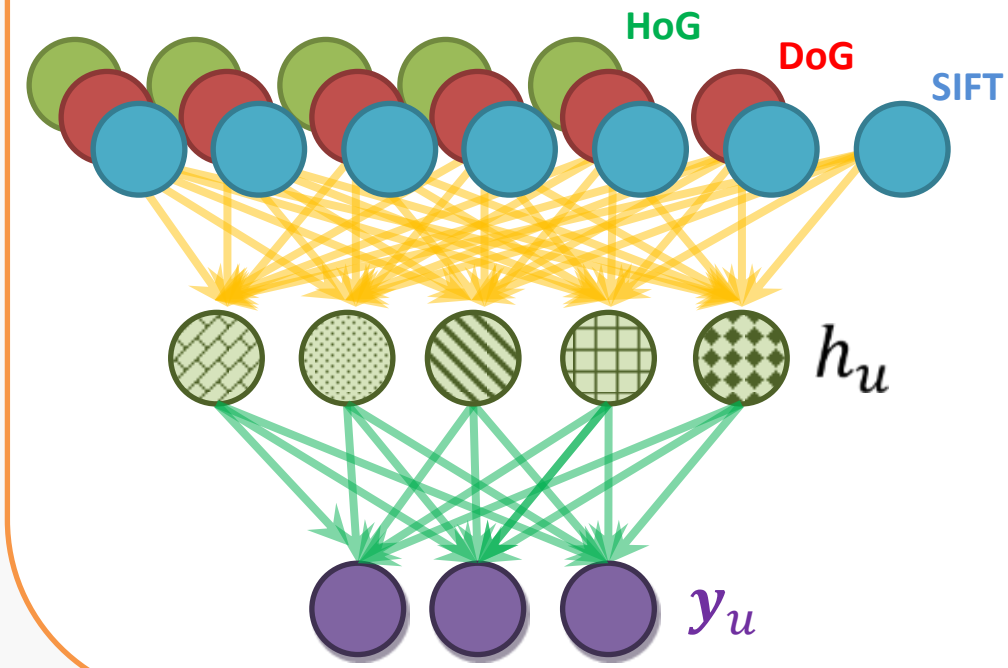
building
car
Road
sky
tree

Embedding Learning → Label Completing

动机： 海量带标签图像，但是标签并不完全；图像可由许多种高维视觉特征表示

算法思路

首先将多种高维视觉特征联合映射到低维嵌入空间中，然后再映射到最终语义空间，解决了联合使用多种不同视觉特征的计算图像间相似度问题，还利用了标签之间的相关性。



Wei Zhang, Ke Zhang, Pan Gu, Xiangyang Xue, Multi-View Embedding Learning for Incompletely Labeled Data, IJCAI 2013

图像标注：在线推荐

提出Hybrid Probabilistic Model --- HPM

❶
$$P(w|I,\mathcal{U}) = \frac{P(I|w)P(\mathcal{U}|w,I)P(w)}{P(I)P(\mathcal{U}|I)}$$

❷
$$P(w|I,\mathcal{U}) = \frac{P(I|w)P(\mathcal{U}|w)P(w)}{P(I)P(\mathcal{U}|I)}$$

❸
$$\log P(w|I,\mathcal{U}) = \log P(I|w) + \sum_{t=1}^{|\mathcal{U}|} \log P$$
$$+ \log P(w) - \log P(I)$$

测试结果	
Manual Annotation	field horses mare foals
HPM, *Given* 0 Annotation	mare foals horses field meadow
HPM, *Given* 1 Annotation	mare foals horses field grass
HPM, *Given* 2 Annotation	field mare foals horses grass

Ning Zhou, William K. Cheung, Guoping Qiu, and Xiangyang Xue. A Hybrid Probabilistic Model for Unified Collaborative and Content-based Image Tagging. IEEE TPAMI 2011

总结

- 在图像分割与语义标注中，利用感知组织特征、标签之间相关性、标签与视觉特征之间相关性等，可有效提升图像语义分析的性能。

- 互联网上已有数千亿幅用户上传图片，其中很多图片还有用户提供的语义标签（Tags），这些大数据（big data？）给**图像语义分析**带来了全新研究挑战：

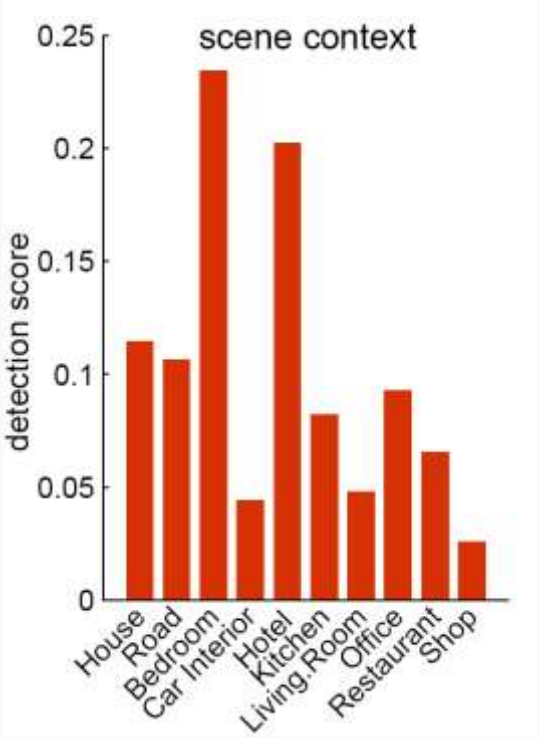
  ✓ **超大规模图像数据**：数千亿幅图像，世间万物皆有图像

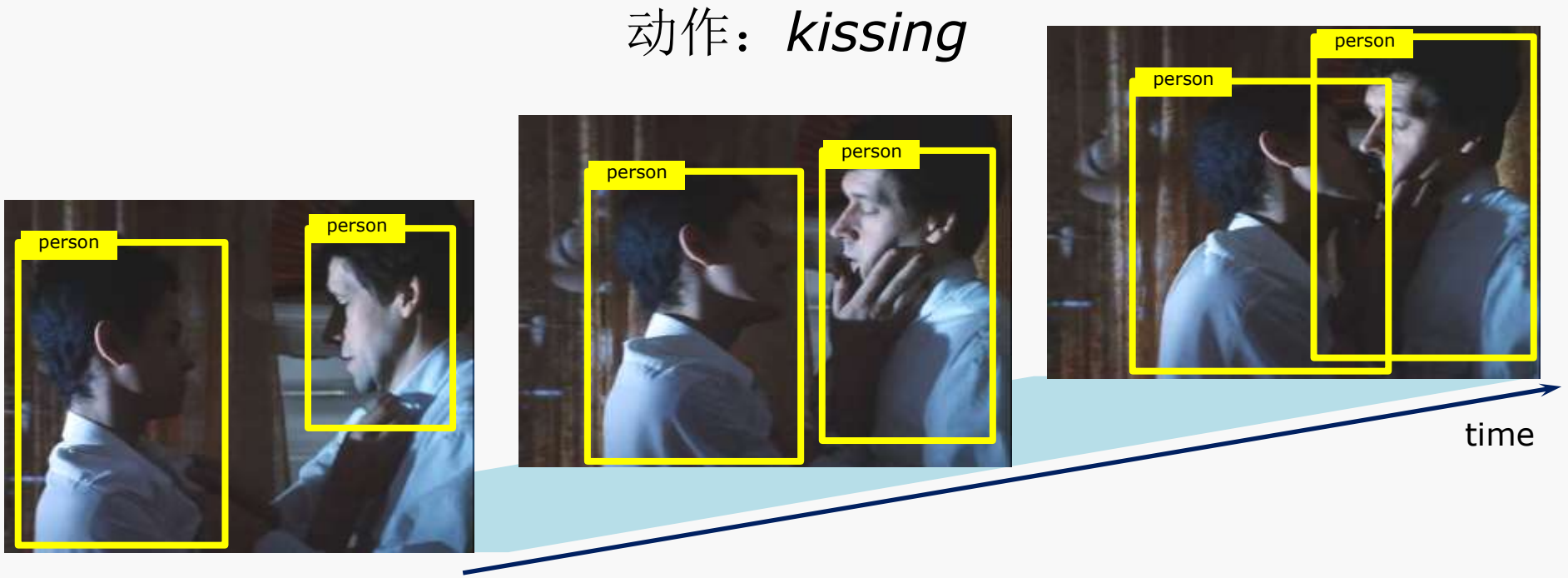
  ✓ **成千上万语义标签**：标签虽多，但是往往不精确、不完全、主观随意…

  ✓ **丰富多彩应用需求**：图像应用层出不穷，对语义分析提出不同需求

视频动作识别：利用上下文信息提升动作识别精度

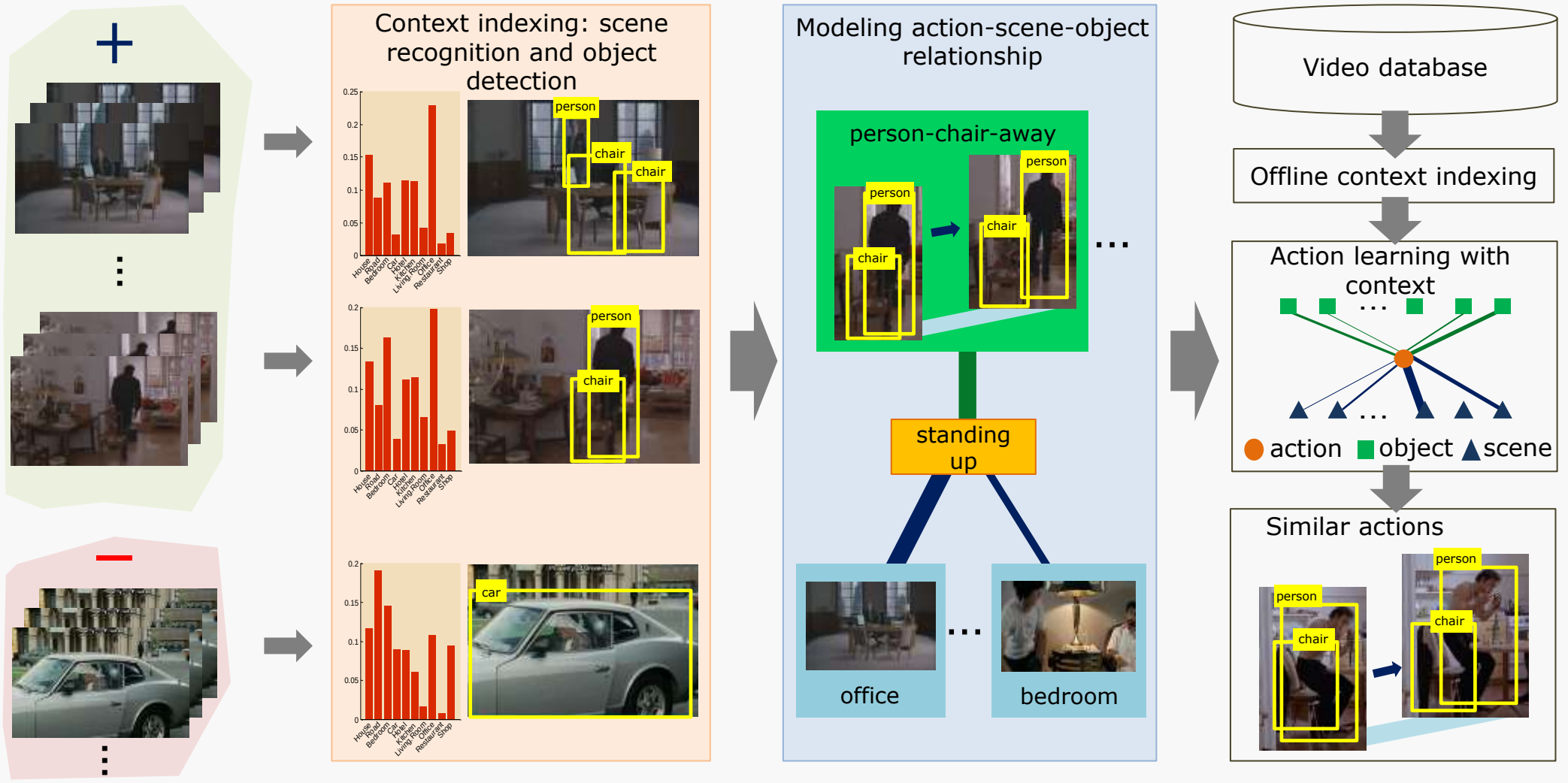
思路：检测运动目标、计算运动目标之间相对运动关系、并考虑所在场景上下文



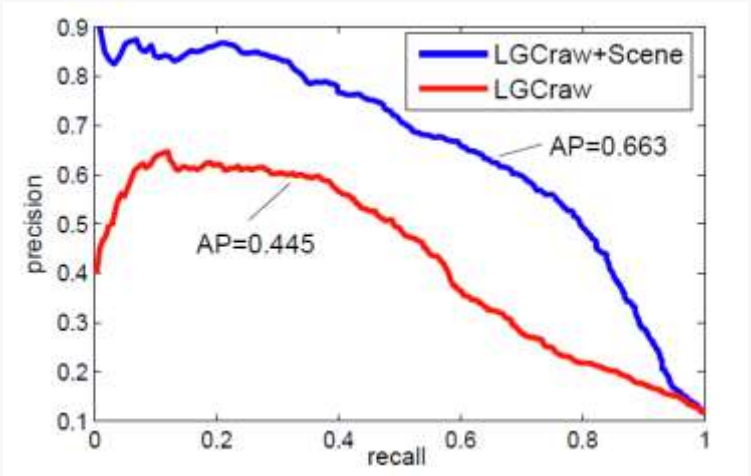


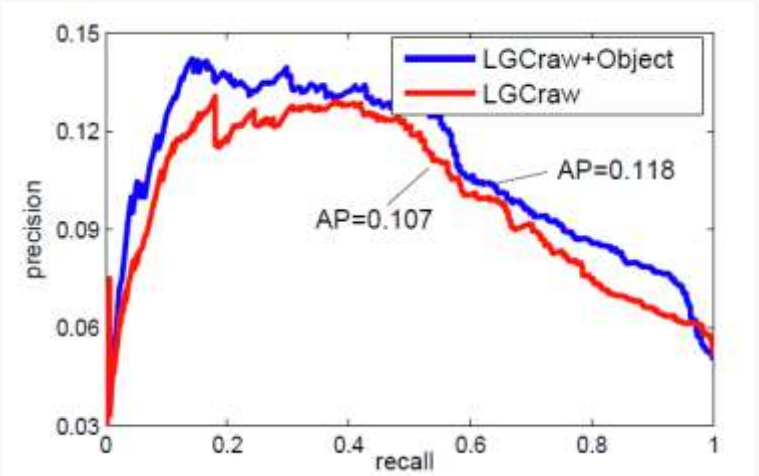
Yu-Gang Jiang, Qi Dai, Jun Wang, Chong-Wah Ngo, Xiangyang Xue, Shih-Fu Chang: Fast Semantic Diffusion for Large-Scale Context-Based Image and Video Annotation. IEEE Transactions on Image Processing 21(6): 3080-3091 (2012)

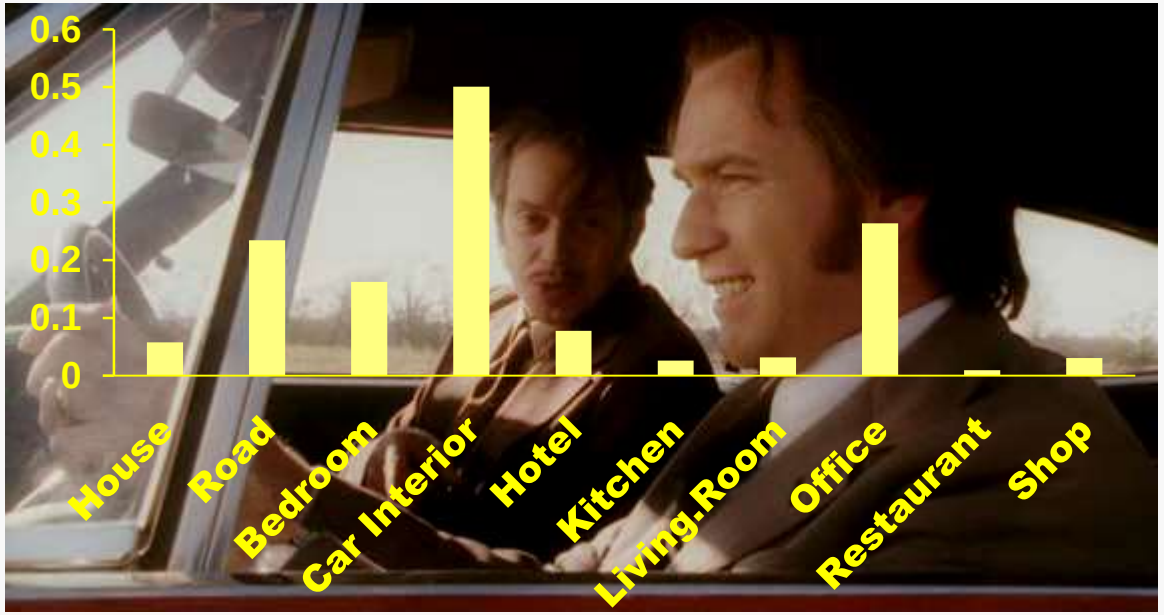
视频动作识别：算法框架


+
−
Context indexing: scene recognition and object detection
person
chair
chair
person
chair
car
Modeling action-scene-object relationship
person-chair-away
person
chair
person
chair
...
standing up
office
...
bedroom
Video database
Offline context indexing
Action learning with context
action
object
scene
Similar actions
person
chair
person
chair

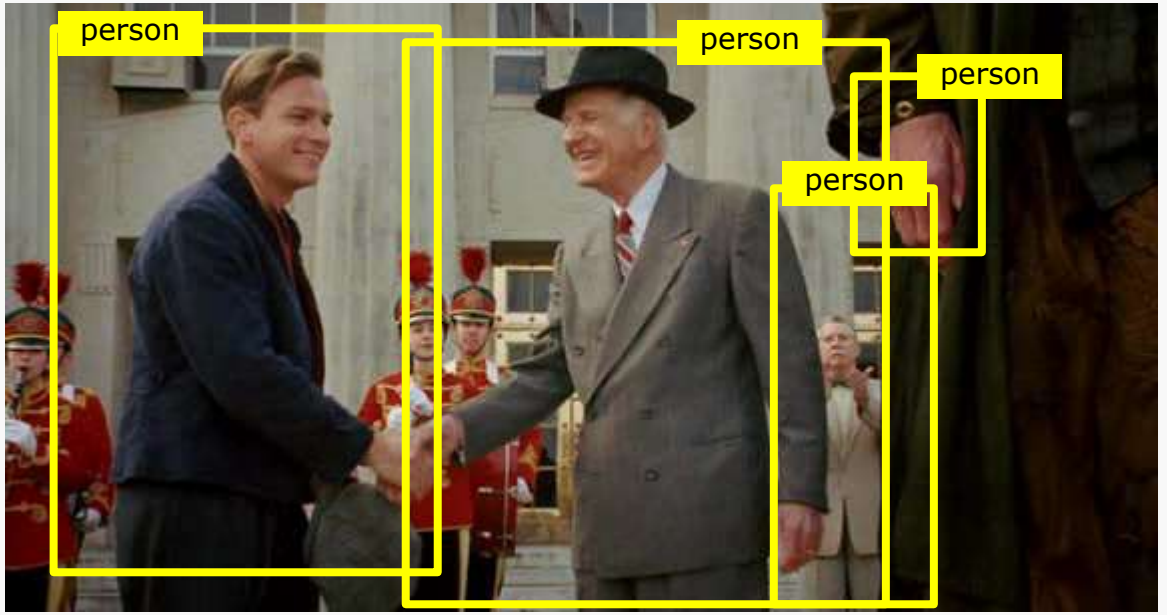
视频动作识别：实验结果







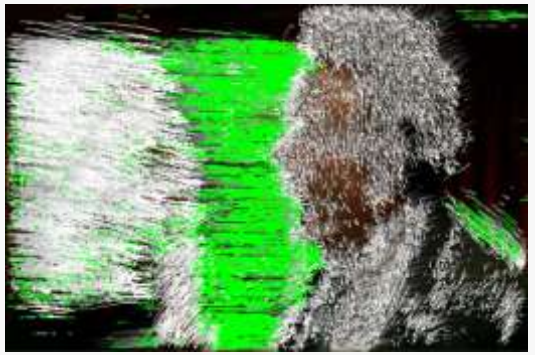
driving car



shaking hands

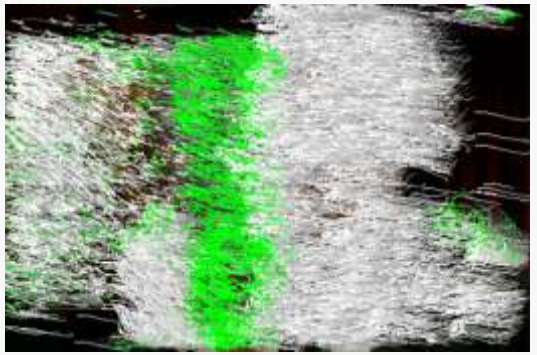
视频动作识别：提出基于轨迹的特征

(a)



(b)

(d)



(c)

(e)

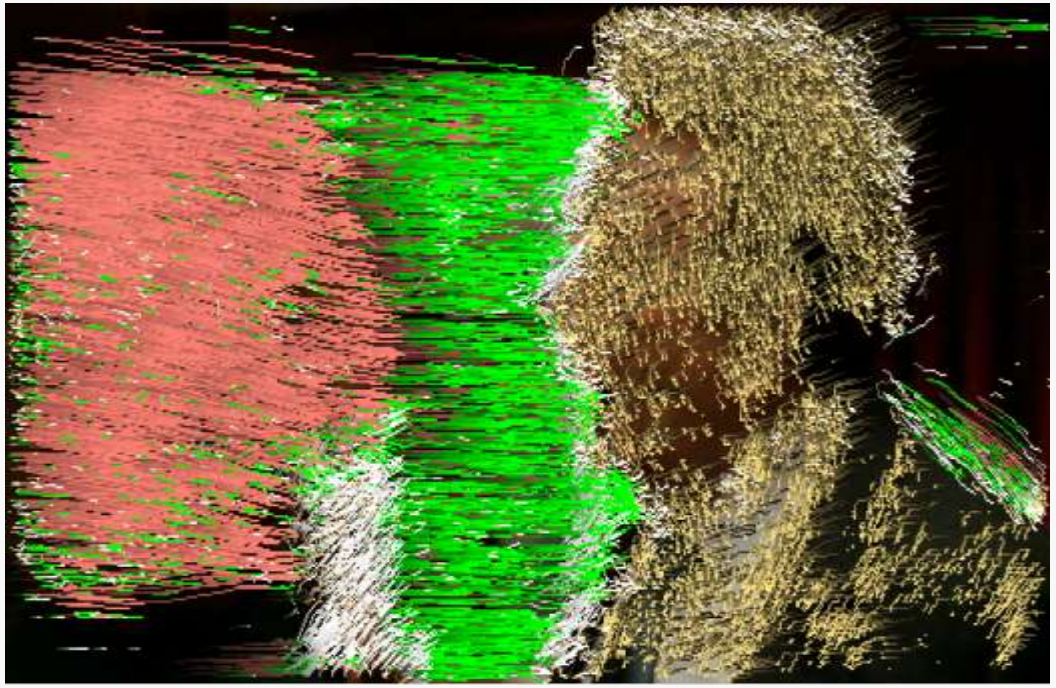
优点：

- 克服相机运动影响
- 捕获运动目标及背景之间关系
- 不需要前景-背景分离
- 容易计算

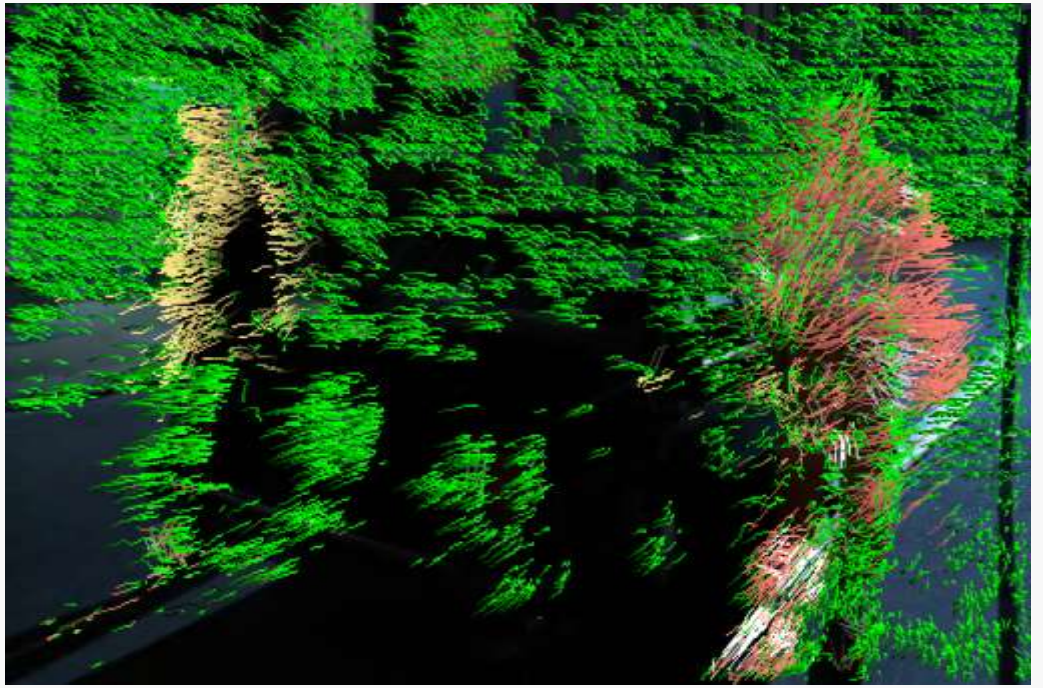
(a) A video frame of a *kissing action*
(b) Local patch trajectories, with the largest trajectory cluster shown in green
(c) Amended trajectories by using the mean motion of the green cluster as a global reference point(全局参考点)
(d) The original patch trajectories, with a trajectory on a person's head shown in red (circled)
(e) Amended trajectories by using the motion of the red trajectory as a local reference point(局部参考点)

视频动作识别：设计思路

通过对轨迹聚类，寻找全局的运动参考点（即理想情况下的静态背景参考点），进而用该参考点校准全部轨迹的运动



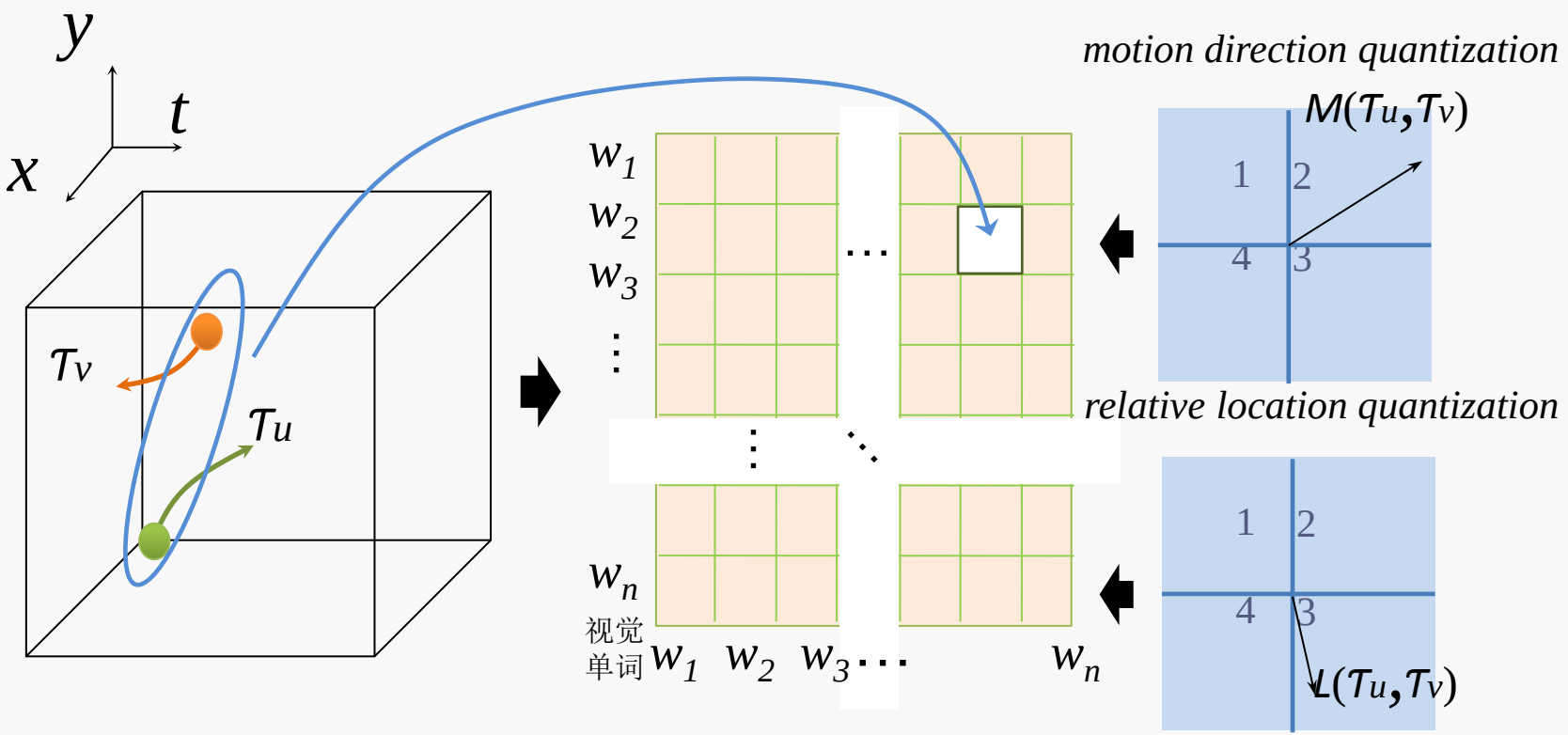
kissing



Two people getting into car

视频动作识别：特征构成

将每一条轨迹当成一个参考点，记录轨迹间的相对运动，彻底解决镜头移动带来的运动评估不准问题



Yu-Gang Jiang, Qi Dai, Xiangyang Xue, Wei Liu, Chong-Wah Ngo: Trajectory-Based Modeling of Human Actions with Motion Reference Points. ECCV (5) 2012: 425-438

视频动作识别：性能测试

Performance of baselines, our representations, and their combination:

	Approach	Hollywood2	Olympic Sports	HMDB51
Baseline results	TrajShape	49.3%	59.5%	24.0%
	4 combined [3]	58.4%	74.3%	37.7%
Our results	TrajShape′	50.2%	59.6%	26.7%
	TrajMF-HOG	39.4%	66.7%	24.3%
	TrajMF-HOF	42.3%	56.0%	25.0%
	TrajMF-MBH	46.9%	74.6%	34.0%
	Our 4 combined	55.6 %	77.6%	39.8%
	All combined	**59.5%**	**80.6%**	**40.7%**

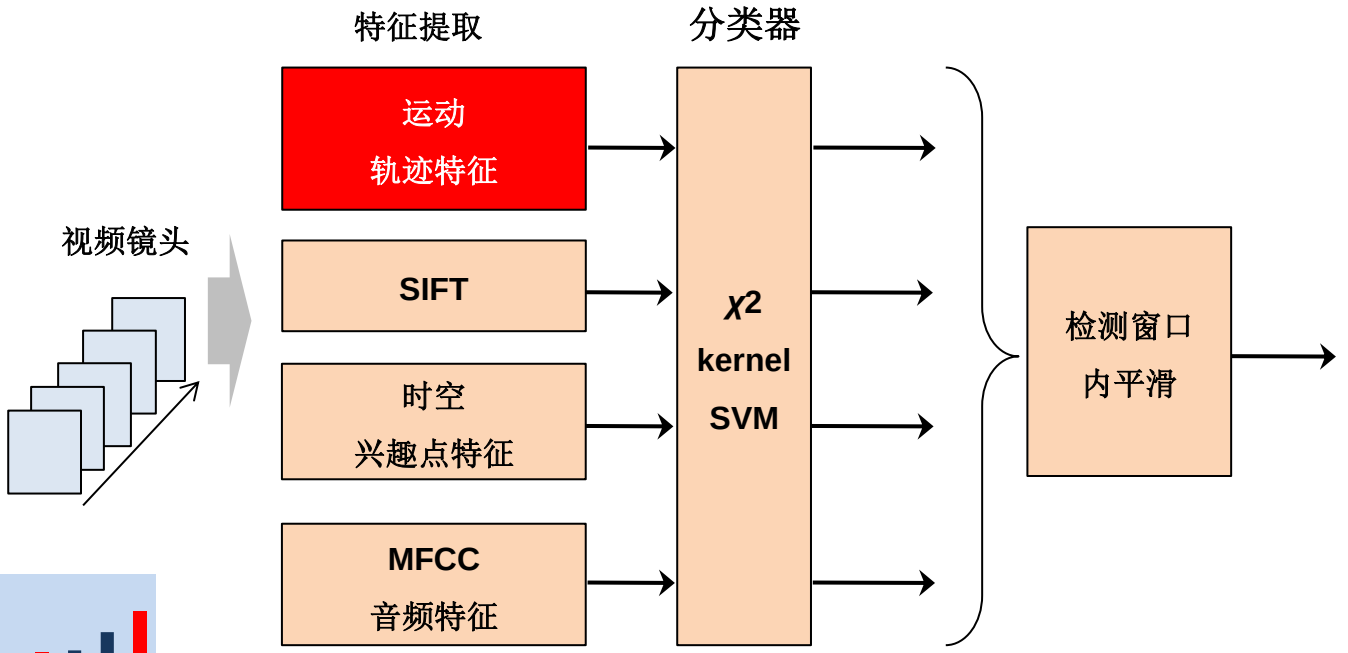
Comparison with the state-of-the-art approaches:

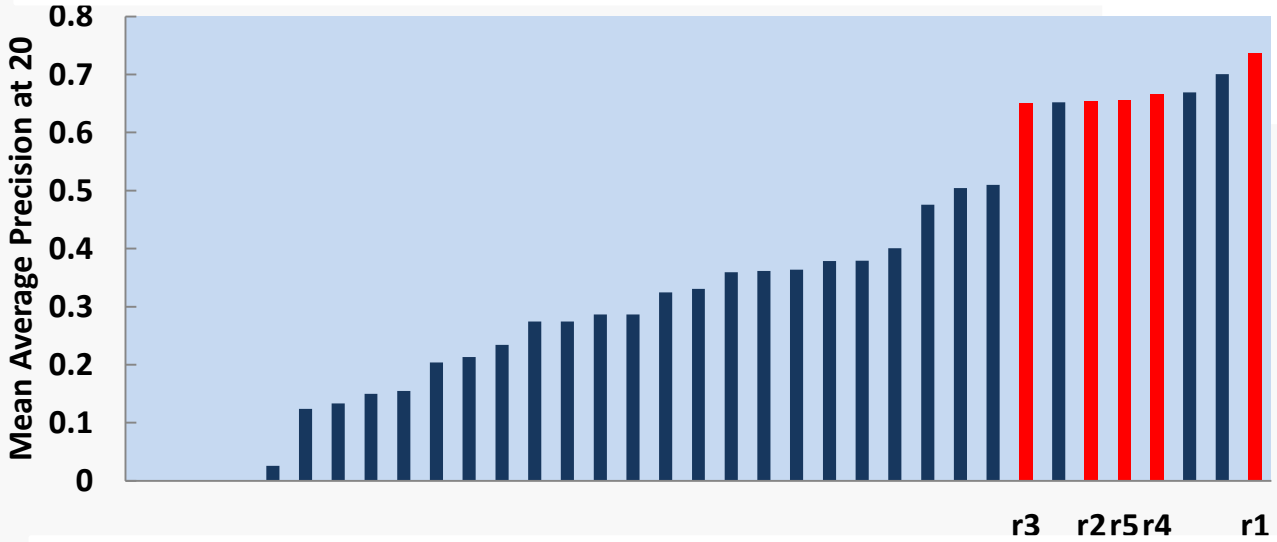
Hollywood2		Olympic Sports		HMDB51	
Taylor et al. [12]	46.6%	Laptev et al. [2]	62.0%	Kuehne et al. [28]	
Gilbert et al. [30]	50.9%	Niebles et al. [33]	72.1%	– HOG-HOF	20.4%
Ullah et al. [34]	53.2%	Liu et al. [35]	74.4%	– C2	22.8%
Le et al. [13]	53.3%	Brendel et al. [36]	77.3%		
Wang et al. [3]	58.3%				
59.5%		**80.6%**		**40.7%**	

Yu-Gang Jiang, Qi Dai, Xiangyang Xue, Wei Liu, Chong-Wah Ngo: Trajectory-Based Modeling of Human Actions with Motion Reference Points. ECCV (5) 2012: 425-438

国际比赛：暴力镜头检测（欧洲MediaEval）

- **目标**：检测视频中的含有暴力行为或事件的片段；暴力的定义是能够造成人体疼痛的动作或事件，如斗殴、爆炸等。
- **应用**：快速生成电影的暴力片段摘要，供家长预览，以决定是否适合儿童观看。





参赛队伍：法国INRIA，英国Imperial College London, 日本NII，复旦（红色）等；2012年和2013年我们都取得了优异成绩！

总结：视频动作识别

- 如何充分挖掘视觉特征（空间域）、运动特征（时间域）、上下文特征（与应用相关先验知识）等，对提高视频动作识别精度的帮助很大！
- 不管是互联网视频，还是监控视频，视频动作识别技术应用前景十分广阔。

近三年发表的相关论文

1. Wei Zhang, Ke Zhang, Pan Gu, Xiangyang Xue, Multi-View Embedding Learning for Incompletely Labeled Data, IJCAI 2013
2. Ke Zhang, Wei Zhang, Yingbin Zheng, Xiangyang Xue, Sparse Reconstruction for Weakly Supervised Semantic Segmentation, IJCAI 2013
3. Yu-Gang Jiang, Yanran Wang, Rui Feng, Xiangyang Xue, Yingbin Zheng, Hanfang Yang: Understanding and Predicting Interestingness of Videos. AAAI 2013
4. Yu-Gang Jiang, Jun Wang, Xiangyang Xue, Shih-Fu Chang: Query-Adaptive Image Search With Hash Codes. IEEE Transactions on Multimedia 15(2): 442-453 (2013)
5. Hong Liu, Hong Lu, Zhaohui Wen, Xiangyang Xue: Gradient Ordinal Signature and Fixed-Point Embedding for Efficient Near-Duplicate Video Detection. IEEE Trans. Circuits Syst. Video Techn. 22(4): 555-566 (2012)
6. Yu-Gang Jiang, Qi Dai, Jun Wang, Chong-Wah Ngo, Xiangyang Xue, Shih-Fu Chang: Fast Semantic Diffusion for Large-Scale Context-Based Image and Video Annotation. IEEE Transactions on Image Processing 21(6): 3080-3091 (2012)
7. Yao Lu, Wei Zhang, Cheng Jin, Xiangyang Xue: Learning attention map from images. CVPR 2012: 1067-1074
8. Yao Lu, Wei Zhang, Hong Lu, Xiangyang Xue: Salient Object Detection using concavity context. ICCV 2011: 233-240
9. Yu-Gang Jiang, Qi Dai, Xiangyang Xue, Wei Liu, Chong-Wah Ngo: Trajectory-Based Modeling of Human Actions with Motion Reference Points. ECCV (5) 2012: 425-438
10. Wei Zhang, Xiangyang Xue, Jianping Fan, Xiaojing Huang, Bin Wu, and Mingjie Liu, Multi-Kernel Multi-Label Learning with Max-Margin Concept Network, IJCAI 2011
11. Xiangyang Xue, Wei Zhang, Jie Zhang, Bin Wu, Jianping Fan, and Yao Lu, Correlative Multi-Label Multi-Instance Image Annotation, ICCV 2011
12. Ning Zhou, William K. Cheung, Guoping Qiu, and Xiangyang Xue. A Hybrid Probabilistic Model for Unified Collaborative and Content-based Image Tagging. IEEE TPAMI 2011

谢谢！

omap.fudan.edu.cn

计算机科学技术学院

复旦大学