

Hierarchical Co-salient Object Detection via Color Names (ACPR 2017)

Jing Lou¹, Fenglei Xu¹, Qingyuan Xia¹
Wankou Yang^{2,*}, Mingwu Ren^{1,*}

¹Nanjing University of Science and Technology ²Southeast University
<http://www.loujing.com/hcn-co-sod/>

(Sep 9, 2017)

Outline

- 1 Introduction
- 2 Related Work
- 3 Color Names Based Hierarchical Model
 - Pipeline
 - Single-Layer Combination
 - Single-Layer Refinement
 - Multi-Layer Fusion and Refinement
 - Color Names Based Co-saliency Detection
- 4 Experiments
 - Data Set
 - Evaluation Metrics
 - Parameter Analysis
 - Evaluation of Saliency Fusion
 - Comparisons with Other Models
- 5 Conclusion

I. Introduction

任务

- 视觉显著性的一个研究分支
- 不同于在单张图像中检测显著物体，协同显著物体检测的任务是从一对图像（即两张图像）或者一组图像中检测出共有的显著区域

本文

- 作为尝试，仅研究面向一对图像的协同显著物体检测
- 自底向上的、数据驱动的方式
- 采用分层结构
- 融合现有的两个显著物体检测模型
- 基于全局的颜色特征对比
- 在单图显著性的基础上，检测图像间的协同显著性

II. Related Work

Boolean Map Based Saliency (BMS)

- J. Zhang and S. Sclaroff, “Saliency detection: A boolean map approach,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 153–160.
- 动机：基于 Gestalt 理论的图形 – 背景分离 (figure-ground segregation), 即去除非环绕 (surroundedness) 的区域
- 缺点：仅利用了环绕性线索，其效果类似于在强度图上进行检测，在一定程度上忽略了颜色信息

II. Related Work

Color Name Space Based Saliency (CNS)

- J. Lou, H. Wang, L. Chen, Q. Xia, W. Zhu, and M. Ren, “[Exploiting color name space for salient object detection](#),” arXiv:1703.08912 [cs.CV], pp. 1–13, 2017.
- 动机：为了克服 BMS 方法的缺陷，我们将 BMS 从 LAB 颜色空间扩展到颜色名空间 (Color Name Space)，并引入两个全局颜色线索与 BMS 的环绕性线索相融合
- 缺点：由于同样基于环绕性线索去除贴边区域，故对于即使是轻微贴边的显著物体，CNS 方法检测失效
- 解决手段：引入 RBD 模型

II. Related Work

Robust Background Detection Based Saliency (RBD)

- W. Zhu, S. Liang, Y. Wei, and J. Sun, “Saliency optimization from robust background detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2814–2821.
- 动机：已有方法中，即使目标仅轻微接触图像边界，仍会导致失败
- 边界连通性 (boundary connectivity)

$$BndCon(R) = \frac{|\{p|p \in R, p \in Bnd\}|}{\sqrt{|\{p|p \in R\}|}}$$

- 背景加权的对比模型 (background weighted contrast)

$$wCtr(p) = \sum_{i=1}^N \underbrace{d_{app}(p, p_i)}_{\text{边的权重}} \underbrace{w_{spa}(p, p_i)}_{\text{距离相关}} w_i^{bg}$$

III. Color Names Based Hierarchical Model

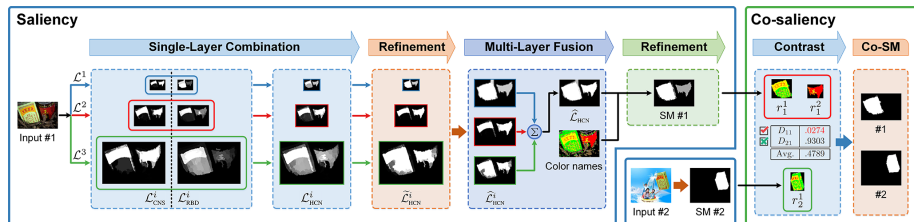


Figure 1: HCN 流程图

优点：可按照流程，分模块调试和改进算法

- (1) 每张图像在 3 个尺度上分别计算单层显著图 (single-layer saliency map)，包括单层融合与改进 (single-layer combination and refinement) 两个步骤
- (2) 融合 3 层结果得到单图显著图 (single-image saliency map，简记为 SM)，包括多层融合与改进 (multi-layer fusion and refinement) 两个步骤
- (3) 在一对图像之间，利用已有的显著区域进行颜色对比，去除非协同显著区域，得到各自的协同显著图 (co-saliency map，简记为 Co-SM)

A. Single-Layer Combination

构建分层结构

考虑到对显著物体大小变化的适应性，构建 3 层结构 ($\mathcal{L}^1, \mathcal{L}^2, \mathcal{L}^3$)，每层的图像宽度分别固定为 100px, 200px, 400px

在单层中融合 CNS 和 RBD 模型

对于第 $\mathcal{L}^i$ 层，分别使用 CNS 和 RBD 模型得到显著图 $\mathcal{L}_{\text{CNS}}^i$ 和 $\mathcal{L}_{\text{RBD}}^i$ ，再融合两者结果得到单层显著图 $\mathcal{L}_{\text{HCN}}^i$ ，其中坐标 (x, y) 处的显著值为：

$$\begin{aligned} \mathcal{L}_{\text{HCN}}^i(x, y) = & (w_f \mathcal{L}_{\text{CNS}}^i(x, y) + (1 - w_f) \mathcal{L}_{\text{RBD}}^i(x, y)) \\ & \times \underbrace{\left(2e^{-|\mathcal{L}_{\text{CNS}}^i(x, y) - \mathcal{L}_{\text{RBD}}^i(x, y)|} - 2e^{-1} + 1 \right)}_{\text{显著性一致性 (consistency)}} \end{aligned} \quad (1)$$

上式中 consistency 项的直观解释：倾向于两个融合模型产生相似的显著性结果，如是则 $\mathcal{L}_{\text{HCN}}^i(x, y)$ 将获得较大的显著值；如果两者截然不同，则显著值趋向于 0 (后面的“+1”项用来避免此情况下得不到任何显著区域)

A. Single-Layer Combination

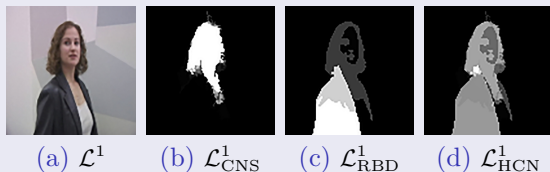


Figure 2: 单层融合结果示例

注：尝试了 48 种不同的底层融合算法 (包括简单地将两者平均、构造中心权重矩阵等)，并使用 3 层 MaxF 指标的均值评价性能，最终确定如式 (1) 的方式

边框裁剪 (Border Effect)

- 部分图像带有边框，将影响 CNS 的输出，故基于边缘图进行裁边操作
- 在构建层次结构之前，先执行图像裁边 (环境：MATLAB R2017a)
- 对于 RBD，裁边后再将 3 层分别送入其超像素分割模块

B. Single-Layer Refinement

动机

- 显著物体检测本质上是一个图像-背景分离问题，故理想的输出是一个二值掩码图像，即显著物体区域具有均匀的显著值 1、背景为 0
- 然而，之前大部分的显著性方法并没有朝这个方向努力
- 在多层融合之前，进一步改进单层显著图，以希望得到 3 张显著区域类似的单层显著图
- 解决手段：由于各层具有相同的颜色分布，考虑使用基于颜色名的全局颜色对比模型来改进单层融合结果 (color names based refinement)
- 注：基于底层图像处理技术，我们尝试了 16 种不同的单层改进算法 (包括不同的形态学处理、显著图校正等)，同样利用 3 层 MaxF 指标的均值来评价各算法性能

B. Single-Layer Refinement

Color Names Based Consistency

- 扩展式 (1) 为

$$J^i = \underbrace{(W^i \circ (\mathcal{L}_{\text{CNS}}^i)^{\circ 2}) \circ (W^i \circ (\mathcal{L}_{\text{RBD}}^i)^{\circ 2})}_{\text{颜色名一致性}} + (\mathcal{L}_{\text{HCN}}^i)^{\circ 2} \quad (2)$$

- 上式中的 W^i 是一个权值矩阵, 通过将 $\mathcal{L}^i$ 转换为颜色名图像 (color name image) 后计算。假设像素 $\mathcal{L}^i(x, y)$ 属于第 k 个颜色名, 则

$$W^i(x, y) = \sum_{j=1}^{11} f_j \|c_k - c_j\|_2^2 \quad (3)$$

f_j 是第 j 个颜色名的频率, c_k 和 c_j 分别是第 k 和第 j 个颜色名的 RGB 值

- 将 $\mathcal{L}_{\text{CNS}}^i \circ \mathcal{L}_{\text{RBD}}^i$ 简记为 C^i , 则式 (2) 中的 J^i 可改写为

$$J^i = (W^i \circ C^i)^{\circ 2} + (\mathcal{L}_{\text{HCN}}^i)^{\circ 2} \quad (4)$$

B. Single-Layer Refinement

后处理步骤 (post-processing)

Algorithm 1 refinement for the saliency map $\mathcal{L}_{\text{HCN}}^i$

Input: C^i and J^i

Output: refined saliency map $\tilde{\mathcal{L}}_{\text{HCN}}^i$

- 1: $\tilde{\mathcal{L}}_{\text{HCN}}^i = \text{RECONSTRUCT}(C^i, J^i)$
- 2: $\tilde{\mathcal{L}}_{\text{HCN}}^i = (\tilde{\mathcal{L}}_{\text{HCN}}^i)^{\circ 2}$ ▷ background suppression
- 3: $\tilde{\mathcal{L}}_{\text{HCN}}^i = \text{ADJUST}(\tilde{\mathcal{L}}_{\text{HCN}}^i, t_a)$ ▷ foreground highlighting
- 4: $\tilde{\mathcal{L}}_{\text{HCN}}^i = \text{HOLE-FILL}(\tilde{\mathcal{L}}_{\text{HCN}}^i)$
- 5: $\mathcal{L}_{\text{HCN}}^i = \text{NORMALIZE}(\tilde{\mathcal{L}}_{\text{HCN}}^i)$

单层改进的优势

- 使用 Hadamard 幂 “ $\circ 2$ ” 抑制背景像素的显著值
- 利用全局颜色对比机制，进一步强调两个融合模型间共有的显著区域
- 利用亮度变换 (ADJUST) 均匀地凸显前景像素， t_a 为 $\tilde{\mathcal{L}}_{\text{HCN}}^i$ 的均值

B. Single-Layer Refinement

单层改进结果示例



Figure 3: (a) CN: 图 2(a)的颜色名图像

- 与单层改进前的结果 (图 2(d)) 相比, 单层改进后的结果 (图 3(e)) 可均匀地凸显出显著区域

C. Multi-Layer Fusion and Refinement

多层融合与改进 (Multi-Layer Fusion and Refinement)

- 与单层处理方式类似，包括两步处理：多层融合和融合结果改进
- 预处理：在多层融合之前，将 3 层的显著图 ($\tilde{\mathcal{L}}_{\text{HCN}}^i$, $i = 1, 2, 3$) 统一缩放到边框裁剪后的图像大小 (详见 Border Effect)。如果在边框裁剪阶段，某图被判为包含边框，则将边框补全，并将边框内像素的显著值设为 0
- 注：我们尝试了 6 种不同的多层处理算法 (主要是针对权重系数的选择和不同的加权方式)，同样采用 MaxF 指标评价不同处理算法的性能

C. Multi-Layer Fusion and Refinement

多层融合 (Multi-Layer Fusion)

- 基于各层与 3 层平均显著图的偏差值 d_i , 计算加权的多层融合显著图。定义 3 层平均显著图 $\bar{\mathcal{L}}_{\text{HCN}} = \frac{1}{3} \sum_{i=1}^3 \hat{\mathcal{L}}_{\text{HCN}}^i$, 则 d_i 为

$$d_i = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N |\bar{\mathcal{L}}_{\text{HCN}}(x, y) - \hat{\mathcal{L}}_{\text{HCN}}^i(x, y)| \quad (5)$$

- 多层融合结果为基于 d_i 的多层加权, 计算公式为

$$\hat{\mathcal{L}}_{\text{HCN}} = \sum_{i=1}^3 \left(\exp\left(-\frac{d_i}{2\bar{d}}\right) \cdot \hat{\mathcal{L}}_{\text{HCN}}^i \right) \quad (6)$$

其中 $\bar{d} = \sum_{i=1}^3 d_i$

- 这里的多层融合不是简单的各层平均, 利用上式, 多层融合结果 $\hat{\mathcal{L}}_{\text{HCN}}$ 将更倾向于与 $\bar{\mathcal{L}}_{\text{HCN}}$ 相似的单层显著图

C. Multi-Layer Fusion and Refinement

多层融合结果改进 (Refinement)

- 改进算法与第 III-B 节中对于单层改进的方法类似，由于是针对 3 张图，故 Hadamard 幂改为 “ $\circ 3$ ”

$$\hat{C} = \hat{\mathcal{L}}_{\text{HCN}}^1 \circ \hat{\mathcal{L}}_{\text{HCN}}^2 \circ \hat{\mathcal{L}}_{\text{HCN}}^3 \quad (7)$$

$$\hat{J} = (\hat{W} \circ \hat{C})^{\circ 3} + (\hat{\mathcal{L}}_{\text{HCN}})^{\circ 3} \quad (8)$$

$$S_s = (\text{RECONSTRUCT}(\hat{C}, \hat{J}))^{\circ 3} \quad (9)$$

其中 $\hat{W}$ 是原图的权值矩阵，计算方式与第 III-B 节中的 W^i 相同。

- S_s 为单图显著图，即所提出的 HCN 方法的显著性检测结果

C. Multi-Layer Fusion and Refinement

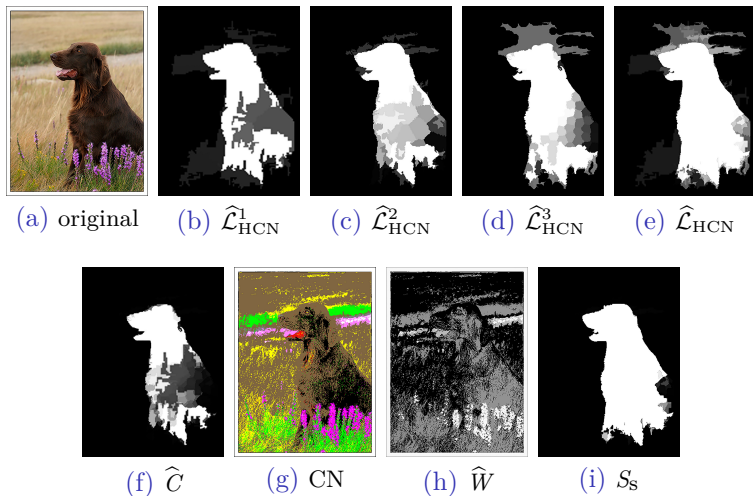


Figure 4: 多层融合与改进示例

与 (b)-(d) 单层结果和 (e) 多层融合结果相比，单图显著图 S_s 更为精确

D. Color Names Based Co-saliency Detection

协同显著性检测原则

- 基于单图显著图，寻找在多图中重复的特征线索

$$\text{Co-saliency} = \text{Saliency} \times \text{Repeatedness} \quad (10)$$

- 本文基于颜色对比特征，该颜色特征同样衍生自颜色名 (Color Names)
- 动机：不同图像间的协同显著区域，具有高度的颜色相似性

D. Color Names Based Co-saliency Detection

- (1) 使用自适应阈值 $\frac{2}{W \times H} \sum_{x=1}^W \sum_{y=1}^H S_s(x, y)$ 分割单图显著图
- (2) 利用得到的各连通区域 r ，在原图中抠出对应的子图，并基于 Color Names 计算该区域的平均颜色 $A(r) = \sum_{j=1}^{11} f_j c_j$
- (3) 对于一对图像中的两张图像，分别两两计算不同区域间的平均颜色差异

$$D_{ij} = \text{Diff}(r_1^i, r_2^j) = \left\| A(r_1^i) - A(r_2^j) \right\|_2^2 \quad (11)$$

其中， r 的下标 1 或 2 表示该区域所在的图像编号，上标 i 或 j 表示该区域在自身图像内的 ID 号

- (4) 计算所有区域对 (r_1^i, r_2^j) 的平均颜色差异 D_{ij} 的均值 $\bar{D}$ ，如果某对区域的平均颜色差异值大于该均值，则该对区域都被视为非协同显著区域，并从各自所在的单图显著图中去除（注：当某个区域同时参与多次比较时，只要有一次被判定为协同显著区域，则予以保留）
- (5) 单图显著图中剩下的显著部分构成最终的协同显著图

IV. Experiments

A. Data set

- Image Pair 数据集包括 150 对图像，即 210 张图像
- 人工标注的显著区域面积大小，平均占所在图像面积大小的 23.87%

B. Evaluation Metrics

$$\begin{aligned} \text{Precision} &= \frac{|M \cap G|}{|M|}, \quad \text{Recall} = \frac{|M \cap G|}{|G|}, \\ F_\beta &= \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}} \end{aligned} \quad (12)$$

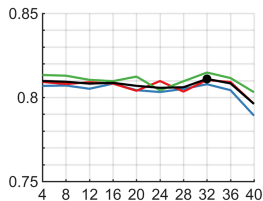
- 固定阈值分割 ($T_f \in [0, 255]$): 除绘制精度-召回曲线和 F_β 曲线外，也比较了 AvgF 和 MaxF
- 自适应阈值分割 (T_a 为 2 倍均值): 除绘制精度-召回柱状图外，也比较了 AdaptF

C. Parameter Analysis

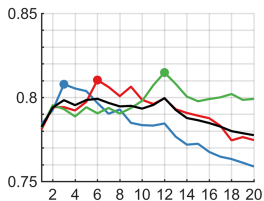
参数分析

- HCN 方法包括 6 个参数：采样间隔 δ ，形态学处理中两个结构元素的核半径 ω_c 和 ω_r ，亮度变换中的饱和系数 ϑ_r 和 gamma 系数 ϑ_g ，以及一个额外的权重系数 w_f 控制单层的显著图融合
- 前 5 个参数的取值范围与 CNS 方法中的相同， w_f 的取值范围被经验地设置为 $[0.1 : 0.1 : 0.9]$
- 考虑到分层结构中各层所具有的尺度大小不同，两个核半径 ω_c 和 ω_r 在每层上分别执行调参操作；而对参数 δ 、 ϑ_r 和 ϑ_g ，则在所有层上取相同的参数值
- 调参依据为寻找 MaxF 曲线的峰值点
- 除 ω_c 和 ω_r 外，其它参数调参是通过 3 层 MaxF 的均值曲线（见参数曲线图中的黑线）上寻找峰值点

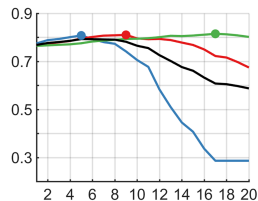
C. Parameter Analysis



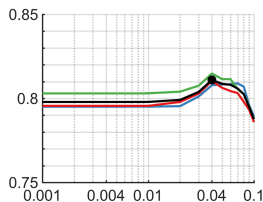
(a) δ



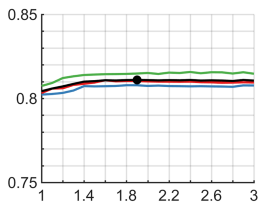
(b) ω_c



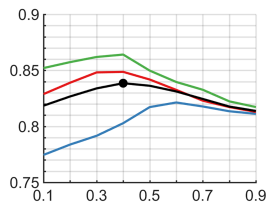
(c) ω_r



(d) ϑ_τ



(e) ϑ_g



(f) w_f

D. Evaluation of Saliency Fusion

融合算法的有效性评估

- 分别评估了单层融合 ($\mathcal{L}_{\text{HCN}}^i$)、单层改进 ($\tilde{\mathcal{L}}_{\text{HCN}}^i$)、多层融合与改进后 (S_s) 的算法性能
- 结果在一定程度上表明了相关融合和改进算法的有效性

Table I: MaxF statistics of single-layer combination

Model	$i = 1$	$i = 2$	$i = 3$	Average
$\mathcal{L}_{\text{CNS}}^i$	.8078	.8103	.8148	.8110
$\mathcal{L}_{\text{RBD}}^i$	.7597	.8288	.8526	.8137
$\mathcal{L}_{\text{HCN}}^i$	.8029	.8487	.8641	.8386
$\tilde{\mathcal{L}}_{\text{HCN}}^i$	.8148	.8510	.8657	.8438

Table II: F_β statistics of multi-layer fusion

Layer	AvgF	MaxF	AdaptF	Average
$\hat{\mathcal{L}}_{\text{HCN}}^1$	.7995	.8148	.8027	.8056
$\hat{\mathcal{L}}_{\text{HCN}}^2$	.8391	.8510	.8435	.8445
$\hat{\mathcal{L}}_{\text{HCN}}^3$	.8568	.8657	.8591	.8605
S_s	.8587	.8663	.8611	.8621

E. Comparisons with Other Models

与其它方法的比较

比较了 14 个显著性方法和 6 个协同显著性方法，比较时分别用 HCN_s 和 HCN_{co} 以示区分

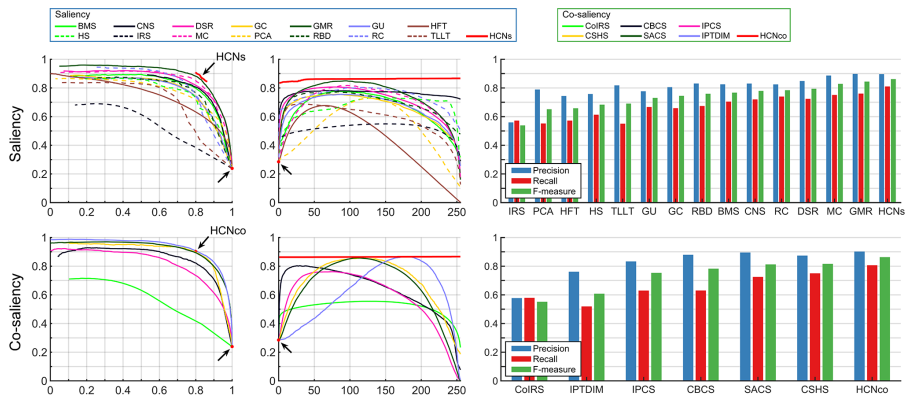


Figure 5: 结果曲线和柱状图

E. Comparisons with Other Models

说明

- 上图中的曲线图通常看起来会觉得比较怪异，但是：
- 精度-召回曲线：由于得到的显著图/协同显著图接近于二值图，故当 $T_f \geq 1$ 时，HCN 的曲线被压缩在一个较小的区间内；当 $T_f = 0$ 时，召回为 1，精度为 0.2387 (即 23.87% 的平均显著区域面积，请见第 IV-A 节)
- F_β 曲线：同样的道理，当 $T_f = 0$ 时， F_β 为 0.2848；而当 $T_f \geq 1$ 时，HCN 的曲线更为平坦，这将便于后续的显著物体分割任务，因为在很宽泛的阈值区间内任意选择一个阈值，都可以得到相似的分割结果
- 回到开始，显著物体检测本质上是一个图形-背景分离问题，而显著图是为显著物体分割所服务。如果显著性检测可以得到一张足够精确的二值显著图，那将大大减少后续分割任务的难度，因为阈值通常并不是那么容易确定。

E. Comparisons with Other Models

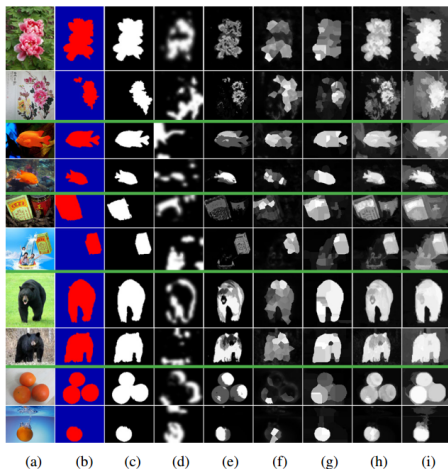


Figure 6: 视觉比较: (a) 输入图像, (b) 人工标注 [13], 以及分别使用 (c) 本文的 HCN 方法, (d) CoIRS [11], (e) CBCS [12], (f) IPCS [13], (g) CSHS [14], (h) SACS [15] 和 (i) IPTDIM [16] 得到的显著图。

E. Comparisons with Other Models

Table III: F_β statistics of saliency models

#	Model	AvgF	MaxF	AdaptF	Average
1	BMS [7]	.6592	.7763	.7666	.7340
2	CNS [17]	.7612	.7817	.7787	.7738
3	DSR [30]	.7098	.8063	.7945	.7702
4	GC [31]	.6634	.7553	.7446	.7211
5	GMR [23]	.7391	.8493	.8442	.8109
6	GU [31]	.6642	.7553	.7303	.7166
7	HFT [8]	.4421	.6772	.6575	.5923
8	HS [3]	.6688	.7345	.6826	.6953
9	IRS [11]	.5149	.5491	.5380	.5340
10	MC [32]	.6933	.8171	.8280	.7795
11	PCA [33]	.5251	.7277	.6506	.6345
12	RBD [26]	.6950	.7727	.7587	.7422
13	RC [19]	.7383	.8031	.7840	.7751
14	TLLT [4]	.5885	.6892	.6908	.6561
15	HCN _s	.8587	.8663	.8611	.8621
Average		.6614	.7574	.7407	.7198

Table IV: F_β statistics of co-saliency models

#	Model	AvgF	MaxF	AdaptF	Average
1	CoIRS [11]	.5150	.5548	.5512	.5403
2	CBCS [12]	.6433	.8028	.7816	.7425
3	IPCS [13]	.5855	.7612	.7526	.6998
4	CSHS [14]	.6894	.8559	.8157	.7870
5	SACS [15]	.6499	.8571	.8114	.7728
6	IPTDIM [16]	.6161	.8671	.6070	.6968
7	HCN _{co}	.8620	.8665	.8625	.8637
Average		.6516	.7951	.7403	.7290

定量比较

- 协同显著检测方法 HCN_{co} 的 MaxF 指标略低于 IPTDIM 方法 (约 0.692‰), 但其它两个指标较之大幅领先, 这也是结果显著图接近于二值图像的优势之一

V. Conclusion

结论

- 基于两个现有的显著性模型和一个颜色命名模型
- 是利用 Color Names 进行显著性检测的另一个案例研究
- Color Names 同样适用于协同显著性检测

未来的工作

- 基于其它手段改善性能，考虑由于引入 Color Names 带来的负面效果 (如仅通过颜色很难与背景分离)
- 改进协同显著性的度量机制，将方法从一对图像扩展到对一组图像的处理

Thanks