

基于主客观联合评价判据的红外超分辨率重建方法

邵保泰^{1,2,3}, 汤心溢^{1,3*}, 张宁

(1. 中国科学院上海技术物理研究所, 上海 200083;

2. 中国科学院大学, 北京 100049;

3. 中国科学院智能红外感知重点实验室, 上海 200083)

摘要: 在红外超分辨率成像应用中, 提高主观视觉效果有着很现实的需求. 当前基于深度学习的图像超分辨率重建方法大多以客观评价指标为损失函数进行训练和优化, 主观评价方法因量化困难而难以应用, 为此本文着重研究了主观评价和各种可量化的客观评价指标的相关性, 发现相位一致性特征与主观评价结果关联度高. 据此设计了基于主客观联合评价的损失函数, 应用于红外图像超分辨率重建算法, 实验表明, 在保持客观质量评价的同时, 更好地提高了图像的主观视觉效果.

关键词: 超分辨率重建; 红外图像; 主客观联合评价; 损失函数

中图分类号: TP391.41 文献标识码: A

Research on infrared super-resolution based on criterion of subjectivity and objectivity joint evaluation

SHAO Bao-Tai^{1,2,3}, TANG Xin-Yi^{1,3*}, ZHANG Ning

(1. Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Key Laboratory of Intelligent Infrared Perception, Chinese Academy of Sciences, Shanghai 200083, China)

Abstract: In the application of infrared super-resolution imaging, it is actually necessary to improve the subjective visual effects of images. At present, most image super-resolution reconstruction methods based on deep learning are trained and optimized with objective evaluation index as loss function. The subjective evaluation methods are difficult to apply due to the difficulty of quantification. Therefore, this paper focuses on the correlation between subjective evaluation and quantifiable objective evaluation indexes, and finds that the characteristics of phase consistency are highly correlated with subjective evaluation results. Based on this, a loss function based on subjective and objective joint evaluation is designed and applied to the super-resolution reconstruction algorithm of infrared image. Experiments show that this method can improve the subjective visual effect of image while maintaining the objective evaluation score.

Key words: super-resolution reconstruction, infrared image, subjective and objective joint evaluation, loss function

PACS: 42. 30. Va, 42. 30. Wb, 07. 05. Pj

引言

图像超分辨率(image super-resolution restora-

tion, SR)重建是指利用一幅或多幅低分辨率(low-resolution, LR)图像恢复成高分辨率(high-resolu-

收稿日期: 2019-04-26, 修回日期: 2019-09-12

Received date: 2019-04-26, Revised date: 2019-09-12

基金项目: 国家十三五国防预研项目(HJJ2017-0083/YYAB3007), 中国科学院青年创新促进会(2014216)

Foundation items: Supported by Thirteen Five national defense research Foundation(HJJ2017-0083/YYAB3007), and Youth Innovation Promotion Association(2014216)

作者简介(Biography): 邵保泰(1993-), 男, 山东聊城人, 博士研究生, 主要研究领域为计算机视觉与深度学习. E-mail: shaobaotai@sina.com

* 通讯作者(Corresponding author): E-mail: gq227@mail.sitp.ac.cn

tion, HR)图像. 随着神经网络在计算机视觉上的广泛应用, 基于神经网络的超分辨率重建方法成为了当前研究热点.

超分辨率图像的评价对于重建算法的评估具有重要意义, 目前使用的评价方法有主观评价方法和客观评价方法^[1]. 当前基于深度学习的图像超分辨率重建方法大多以客观评价指标为损失函数进行训练和优化, 主观评价方法因量化困难而难以应用, 但在很多应用中提高主观视觉效果有重要意义. 为此本文对客观评价算法与主观评价方法的相关性进行了研究, 提出了基于主客观联合评价的损失函数, 并应用于红外超分辨率重建算法中, 取得了良好效果.

1 超分辨率重建图像质量主观、客观评价指标及相关性研究

1.1 图像质量主观、客观评价方法

客观评价方法通过对图像进行统计分析或特征计算以定量化表示图像质量, 常用的客观评价方法有峰值信噪比、结构相似性指数和特征相似性指数等.

(a) 峰值信噪比 (peak signal to noise ratio, PSNR) 通过计算待评价图像和参考图像的均方误差来定义, 其定义表达式为:

$$\text{PSNR} = 20 \log_{10} \frac{2^l - 1}{\sqrt{\text{MSE}}}, \quad (1)$$

$$\text{MSE} = \frac{1}{hw} \sum_{x=1}^h \sum_{y=1}^w (\hat{I}_{x,y} - I_{x,y})^2, \quad (2)$$

其中式(1)中 l 表示图像像素亮度值的比特数, 通常为 8, 此时图像像素亮度的阶数为 256.

(b) 结构相似性指数 (structure similarity image measure, SSIM)^[2] 从亮度、对比度、结构三个方面计算图像的相似性, 是待评价图像和参考图像之间相关相似性的综合表示, 其定义式为:

$$\text{SSIM} = \frac{(2\mu_i \mu_{\hat{i}} + C_1)(2\sigma_{i\hat{i}} + C_2)}{(\mu_i^2 + \mu_{\hat{i}}^2 + C_1)(\sigma_i^2 + \sigma_{\hat{i}}^2 + C_2)}, \quad (3)$$

其中 μ_i 表示参考图像的亮度平均值, $\mu_{\hat{i}}$ 表示待评价图像的亮度平均值, σ_i , $\sigma_{\hat{i}}$ 分别是参考图像和待评价图像的方差, $\sigma_{i\hat{i}}$ 表示两幅图像之间的协方差, C_1 , C_2 是两个常数, 防止分母为零.

(c) 特征相似性指数 (Feature Similarity Index, FSIM)^[3] 是一种结合图像频率域特征和空间域梯度特征的图像评价方法. 其计算方式为:

$$\text{FSIM} = \frac{\sum_{(x,y) \in \Omega} S_L(x,y) \cdot PC_m(x,y)}{\sum_{(x,y) \in \Omega} PC_m(x,y)}, \quad (4)$$

式中 Ω 表示整幅图像, $PC_m(x,y) = \max(PC_l(x,y), PC_r(x,y))$ 用来衡量单个像素的相似性在整体图像相似性中的权重. $S_L(x,y)$ 表示综合相似性, 通过计算相位一致性特征和梯度特征获得.

主观评价方法是指以评价者对图像的主观视觉感受为标准, 评价图像质量的优劣. 根据重建算法的需要, 本文采用以原始高分辨率图像作为参考, 对重建图像质量进行主观评分的方法, 邀请了 15 位专业人士对每张图像进行打分, 以红外图像的视觉效果质量和真实度为打分标准, 统计每张红外图像所有观察人员的分数, 求取平均值并线性映射到 $[0, 5]$ 范围内作为图像的主观评分结果.

1.2 主客观评价相关性实验

本文自建了红外图像数据库, 采用中波红外相机 (视场角 $11^\circ \times 8.8^\circ$, 图像尺寸 640×512) 采集上海市街景图像, 共 1 000 余幅, 图像内容包含行人、车辆、交通标识、绿植和建筑等. 从中选取了 50 张内容差异较大、信息丰富的清晰红外图像作为参考图像, 对其进行下采样后 (下采样因子为 2 和 4), 使用双立方插值、SRCNN 和 SRResnet 等方法提高分辨率, 作为待评价图像.

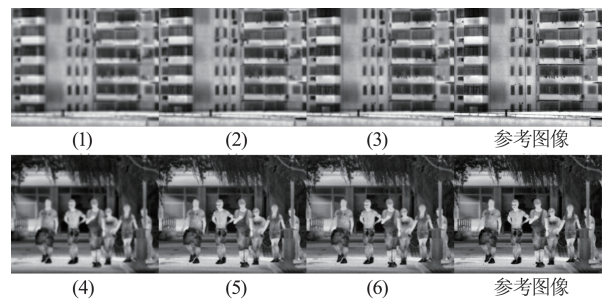


图1 红外图像示例

Fig. 1 Examples of infrared images

待评价图像与参考图像具有相同的分辨率, 使用主、客观评价方法对其进行打分, 获得待评价图像的主观评分结果、峰值信噪比、结构相似性指数和结构相似性指数. 图 1 展示了部分图像示例, 其对应得分如表 1 所示.

为了对比三种客观评价指标与图像主观评分结果的相关性, 本文使用了斯皮尔曼秩相关系数, 肯德尔秩相关系数和皮尔森线性相关系数^[4]. 三种

表1 红外图像质量评价分数

Table 1 Quality evaluation score of infrared image

下采样 因子	超分辨率 方法	主观分 数	PSNR	SSIM	FSIM	
1	4	Bicubic	1.26	23.29	0.624 1	0.546 1
2	4	SRCNN	1.63	25.89	0.686 6	0.626 8
3	4	SRResnet	2.19	25.72	0.728 6	0.700 8
4	2	Bicubic	2.57	29.24	0.848 4	0.752 6
5	2	SRCNN	3.98	30.98	0.909 2	0.887 0
6	2	SRResnet	4.50	30.22	0.926 1	0.921 8

系数的取值范围为 $[-1, 1]$, 系数为正值时表示正相关, 系数为负值时表示负相关, 系数绝对值越大表示相关性越强, 系数为0时表示两个变量不相关。

(a) 斯皮尔曼秩相关系数 (spearman rank order correlation coefficient, SROCC) 根据数据的排序位置计算相关性, 计算之前先对原始数据 (X, Y) 排序, 获得原始数据的位置 (X', Y') , 称为“等级”或“秩”, 然后依据对应数据的等级差计算相关性。

$$SROCC = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}, \quad (5)$$

其中, $d = X' - Y'$ 表示对应数据的等级差, n 表示数据对的数量。

(b) 肯德尔等级相关系数 (Kendall rank order correlation coefficient, KROCC) 也是根据数据的等级来计算。假设有 n 个统计对象, 每个对象均有两个属性, 首先将所有对象按其中一个属性的取值排列, 那么另一个属性取值的排列可能是乱序的。假设所有统计对象中两个属性值排列关系一致的有 P 对, 则有

$$KROCC = \left(P - \left(\frac{n(n-1)}{2} - P \right) \right) / \frac{n(n-1)}{2} = \frac{4P}{n(n-1)} - 1, \quad (6)$$

式中 n 个统计对象中每两个作为一对进行比较, 共有 $\frac{n(n-1)}{2}$ 对。

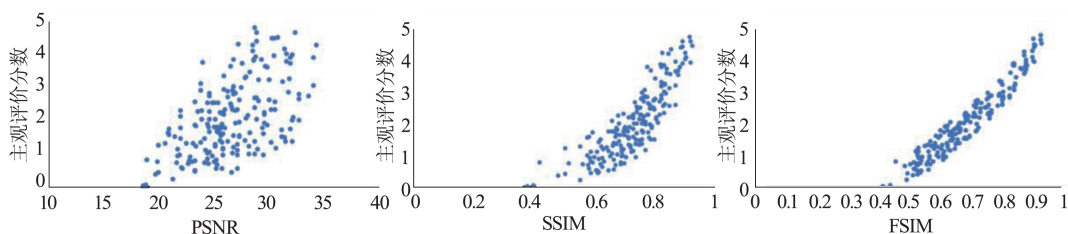


图2 主客观评价分数散点图

Fig. 2 Scatter plots of the subjective and objective evaluation scores

(c) 皮尔森线性相关系数 (Pearson linear correlation coefficient, PLCC) 是一种度量两个变量线性相关程度的统计量, 定义为两个变量的协方差和各自变量标准差乘积的商:

$$PLCC = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}, \quad (7)$$

其中 $\text{cov}(X, Y)$ 表示两个变量 (X, Y) 的协方差, σ_X, σ_Y 分别表示两个变量 (X, Y) 的标准差。

红外图像主、客观评价相关性实验结果如表2所示, 从中看出三种客观评价算法与主观评分结果均为正相关, 其中 PSNR 值与主观评价结果线性相关性最低, FSIM 算法得到的结果与主观评价结果线性相关性最高。图2更直观的表现了客观评价算法与主观评价结果的关系, 从散点图中可以看到 FSIM 与主观评分结果相关性最高, PSNR 值的相关性最低。本文实验结果与文献[3]中可见光实验结果基本一致, 表明在红外图像领域, 特征相似性指数更接近于人类主观视觉感受。

表2 主客观评价相关性结果

Table 2 Results of the correlation between subjective and objective evaluation

	PSNR	SSIM	FSIM
SROCC	0.9072	0.9409	0.9726
KROCC	0.8282	0.8604	0.9004
PLCC	0.8920	0.9327	0.9576

2 主客观联合评价损失函数设计

根据上节研究结果, 本文提出了基于主客观联合评价的损失函数, 结合图像的相位一致性 (phase consistency, PC)^[5] 特征和梯度特征来拟合主观评价结果, 损失函数如式(8)所示:

$$L = \frac{\sum_{\Omega} (\alpha_1 l_{MSE} + \alpha_2 l_{PC} + \alpha_3 l_G) \times PC(I^{HR})}{\sum_{\Omega} PC(I^{HR})}, \quad (8)$$

其中 Ω 表示整幅图像, l_{MSE}, l_{PC}, l_G 表示3个不同的损

失函数分量,计算方式如下:

$$l_{\text{MSE}} = (I^{\text{SR}} - I^{\text{HR}})^2 \quad (9)$$

$$l_{\text{PC}} = (PC(I^{\text{SR}}) - PC(I^{\text{HR}}))^2 \quad (10)$$

$$l_{\text{G}} = (G(I^{\text{SR}}) - G(I^{\text{HR}}))^2 \quad (11)$$

$PC(\cdot)$ 表示图像的相位一致性特征, $G(\cdot)$ 表示图像的梯度特征,本文使用了Sobel算子. $\alpha_1, \alpha_2, \alpha_3$ 表示三种损失分量的比例,本文中 $\alpha_1 = \alpha_2 = \alpha_3 = 1$.

本文采用Log-Gabor小波变换来计算相位一致性^[6].首先使用二维Log-Gabor滤波器对红外图像在方向 θ_j 和尺度 n 下滤波,得到在点 x 处滤波响应向量 $[e_{n,\theta_j}(x), o_{n,\theta_j}(x)]$.然后计算响应向量的幅值和局部能量函数:

$$A_{n,\theta_j}(x) = \sqrt{e_{n,\theta_j}(x)^2 + o_{n,\theta_j}(x)^2} \quad (12)$$

$$I_{n,\theta_j}(x) = \sum_n e_{n,\theta_j}(x), H_{n,\theta_j} = \sum_n o_{n,\theta_j}(x) \quad (13)$$

$$E_{n,\theta_j}(x) = \sqrt{I_{n,\theta_j}^2(x) + H_{n,\theta_j}^2(x)} \quad (14)$$

则红外图像的相位一致性为:

$$PC(x) = \frac{\sum_j E_{n,\theta_j}(x)}{\sum_j \sum_n A_{n,\theta_j}(x) + \varepsilon} \quad (15)$$

3 红外超分辨率重建网络结构

红外图像超分辨率重建网络(IRSR)结构如图3所示,超分辨率网络模型输入的是低分辨率红外图像,输出放大4倍之后的超分辨率红外图像. IRSR可以分为三个部分:红外图像特征提取、特征上采样和红外图像重建.其中图像特征提取结构使用了多尺度稠密残差块;特征上采样结构使用两个像素重组结构将图像特征放大了4倍^[7];最后选用了反卷积层来完成超分辨率红外图像的重建任务^[8].

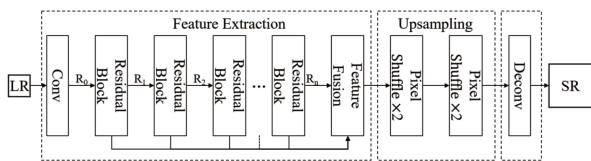


图3 单帧红外图像超分辨率重建网络结构示意图

Fig. 3 Architecture of IRSR

为了检测不同尺度的红外图像特征,本文设计了多尺度的稠密残差块,分别使用了 3×3 和 5×5 两种卷积核的稠密块,如图4所示.每一个稠密块由4个卷积层和3个激活层构成,这些卷积层具有相同数量的卷积核.令 $f_n(n=1, 2, 3)$ 和 $f_{n-1}(n=1,$

$2, 3)$ 表示稠密块第 n 个卷积层的输出和输入,那么第 n 个卷积层的输出为:

$$f_n = \phi(W_n[f_0, f_1, \dots, f_{n-1}] + b_n) \quad (16)$$

其中 W_n, b_n 分别表示第 n 个卷积层的卷积核和偏置项; ϕ 表示激活函数,本文中使用了PReLU函数^[9].

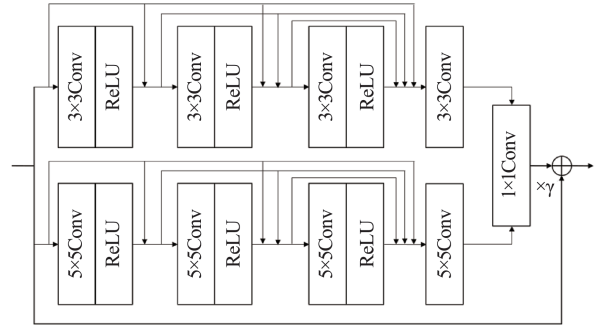


图4 多尺度稠密残差块结构

Fig. 4 Architecture of multi-scale dense residual block

本文使用了 1×1 的卷积层联合了两种尺度的稠密块,同时使用了局部残差学习^[10],假设 $f_{3 \times 3}$ 和 $f_{5 \times 5}$ 表示两个稠密块的输出结果,该操作可以表示为:

$$f_{1 \times 1} = W_{1 \times 1}[f_{3 \times 3}, f_{5 \times 5}] + b_{1 \times 1} \quad (17)$$

$$f = \gamma \times f_{1 \times 1} + f_0 \quad (18)$$

其中 f_0 表示的是稠密块第一个卷积层的输入,也就是残差块的输入, f 表示多尺度稠密残差块的最终输出结果, $\gamma(0 < \gamma < 1)$ 控制残差块的输入与学习到的映射关系的比例,本文 $\gamma = 0.2$.本文网络共使用了9个多尺度稠密残差块,残差块中所有卷积层均使用了64个卷积核,每层输出的特征图通道数均为64.

4 实验结果分析

红外超分辨率重建网络的训练和测试数据集包括了开放数据集和自建数据库,公开数据集为:FLIR数据集(FREE FLIR Thermal Dataset for Algorithm Training)^[11]和CVC数据集(FIR Sequence Pedestrian Dataset)^[12],共14 623张红外图像.在超分辨率重建网络训练过程中,随机选取16幅 96×96 的高清子图像作为小批次进行训练.优化过程选取了 $\beta_1 = 0.9$ 的Adam^[13]算法,首先使用基于均方误差的损失函数^[14]对网络进行训练,学习速率为 10^{-4} ,迭代次数为1 000 000;然后使用本文损失函数对网络进行精调,学习速率为 10^{-5} ,迭代次数为500 000.

实验对比了本文方法和双立方插值、SRCNN^[14]、SRResnet^[1]和SRGAN^[1]等方法的重建结果(图5),图像质量评价如表3所示.使用本文超分

表3 红外超分辨率图像客观评价得分

Table 3 Objective evaluation score of infrared super-resolution images

数据集	评价标准	Bicubic	SRCNN ^[14]	SRRResnet ^[1]	SRGAN ^[1]	Proposed
自建	PSNR	26.15	27.39	28.97	27.89	29.48
	SSIM	0.6549	0.7045	0.7581	0.7849	0.8009
	FSIM	0.5437	0.5911	0.6297	0.6844	0.7516
FLIR	PSNR	28.10	29.66	31.29	31.05	32.10
	SSIM	0.6859	0.7457	0.7895	0.8205	0.8543
	FSIM	0.5553	0.6265	0.7780	0.7834	0.8425
CVC-14	PSNR	21.80	22.82	24.28	23.84	25.21
	SSIM	0.8543	0.8852	0.9315	0.9196	0.9407
	FSIM	0.6461	0.7055	0.7586	0.7809	0.8532

分辨率算法重建出的红外图像质量更接近于原图,物体轮廓更加清晰,纹理细节最为清楚.



图5 红外超分辨率图像(1)

Fig. 5 Infrared super-resolution images (1)

本文进一步实验对比了基于3种不同的损失函数的超分辨率重建效果,损失函数分别为:均方差损失函数,基于判别网络的损失函数^[1]和本文损失函数.

表4展示了三种方法重建红外超分辨率图像的客观评价分数,IRSR-MSE表示网络训练精调阶段使用的是基于均方误差的损失函数,该方法获得了最好的PSNR值,但图像边缘较为模糊,纹理细节不清,如图6所示. IRSR方法使用了本文的损失函数,重建图像更加清晰.同时本文邀请了10位专业人士参考高分辨率图像,对重建图像的视觉效果进行排序.在31幅测试红外图像中,使用主客观联合评

价损失函数的超分辨率图像获得的第一最多,可以认为使用本文方法重建的图像主观视觉效果更好.

表4 红外超分辨率图像评价结果

Table 4 Evaluation results of infrared super-resolution images

数据集	评价标准	IRSR-MSE	IRSR-adv	IRSR(ours)
自建	PSNR	30.19	29.97	29.48
	SSIM	0.7616	0.7990	0.8009
	FSIM	0.7225	0.7202	0.7516
FLIR	PSNR	32.92	32.78	32.10
	SSIM	0.8027	0.8227	0.8543
	FSIM	0.7906	0.8116	0.8425
CVC-14	PSNR	26.90	25.72	25.21
	SSIM	0.9461	0.9507	0.9407
	FSIM	0.8077	0.8391	0.8532
主观评价(排名第一数量)		0	97	213



图6 红外超分辨率图像(2)

Fig. 6 Infrared super-resolution images (2)

5 结论

为了提高超分辨率图像的主观视觉效果,本文研究了图像主客观评价方法相关性,提出了一种使用相位一致性特征和梯度特征的主客观联合损失函数,用来拟合超分辨率图像在主观视觉感受上的评价效果,同时本文利用多尺度稠密残差块设计对

红外图像超分辨率重建网络进行优化改进. 实验结果表明, 和SRCNN等经典超分辨率重建网络相比, 使用本文算法重建的红外超分辨率图像在保持良好的客观评价质量的同时, 具有更好的主观视觉效果.

References

- [1] Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network [J]. *Computer Vision And Pattern Recognition*, 2017: 105-114.
- [2] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13 (4) : 600-612.
- [3] Zhang L, Zhang L, Mou X, et al. FSIM: A feature similarity index for image quality assessment [J]. *IEEE Transactions on Image Processing*, 2011, 20(8) : 2378-2386.
- [4] <https://newonlinecourses.science.psu.edu/stat509/node/155/>
- [5] Morrone M C, Owens R A. Feature detection from local energy [J]. *Pattern Recognition Letters*, 1987, 6 (5) : 303-313.
- [6] Kovési P. Phase congruency: A low-level image invariant [J]. *Psychological Research*, 2000, 64(2) : 136-148.
- [7] Shi W, Caballero J, Huszar F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [J]. *Computer Vision and Pattern Recognition*, 2016: 1874-1883.
- [8] Dong C, Loy C C, Tang X, et al. Accelerating the super-resolution convolutional neural network [J]. *European Conference on Computer Vision*, 2016: 391-407.
- [9] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification [C]. *Proceedings of the IEEE International Conference on Computer Vision*. 2015: 1026-1034.
- [10] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [J]. *Computer Vision and Pattern Recognition*, 2016: 770-778.
- [11] <https://www.flir.com/oem/adas/adas-dataset-form/>
- [12] González A, Fang Z J, Socarras Y, et al. Pedestrian detection at day/night time with visible and fir cameras: a comparison [J]. *Sensors*, 2016, 16(6) : 820.
- [13] Kingma D P, Ba J. Adam: A method for stochastic optimization [J]. *Computer Science*, 2014.
- [14] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2) : 295-307.