



(12)发明专利

(10)授权公告号 CN 106203428 B

(45)授权公告日 2019.04.26

(21)申请号 201610526947.5

(22)申请日 2016.07.05

(65)同一申请的已公布的文献号  
申请公布号 CN 106203428 A

(43)申请公布日 2016.12.07

(73)专利权人 武汉大学  
地址 430072 湖北省武汉市武昌区珞珈山  
武汉大学

(72)发明人 陈震中 丁晓颖

(74)专利代理机构 武汉科皓知识产权代理事务  
所(特殊普通合伙) 42222

代理人 严彦

(51)Int.Cl.

G06K 9/32(2006.01)

G06K 9/46(2006.01)

(56)对比文件

CN 103996198 A,2014.08.20,

CN 101916379 A,2010.12.15,

CN 102693426 A,2012.09.26,

CN 102063623 A,2011.05.18,

US 2012328161 A1,2012.12.27,

审查员 刘海艳

权利要求书2页 说明书5页

(54)发明名称

基于模糊估计融合的图像显著性检测方法

(57)摘要

本发明提供一种基于模糊估计融合的图像显著性检测方法,包括低层视觉特征获取阶段与模糊度特征应用阶段,所述低层视觉特征获取阶段输入待检测的图像,利用经典的自底向上的显著性检测算法对图像进行检测,并将检测得到的显著性特征图作为待检测图像的低层视觉特征获取结果。所述模糊度特征应用阶段,首先输入多组训练图像块,训练稀疏字典;随后利用模糊估计的方式模拟摄影师拍摄时的心理特征并对其进行量化,将量化结果用于指导不同机制下的显著性信息的融合,获取最终的图像显著性检测结果,以提高显著性检测精度。本发明所得到的图像显著性检测结果更加符合人类视觉显著性检测模式。同时,具有更好的鲁棒性与更高的检测精度。

1. 一种基于模糊估计融合的图像显著性检测方法,其特征在于:包括低层视觉特征获取阶段与模糊度特征应用阶段;

所述低层视觉特征获取阶段包含以下步骤,

步骤1.1,输入待检测的测试图像,

步骤1.2,根据步骤1.1输入的测试图像,利用自底向上的图像显著性检测算法获取图像的低层视觉特征,生成待检测图像对应的低层视觉的显著性特征图,用 $S_B$ 表示;

所述模糊度特征应用阶段包含以下步骤:

步骤2.1,输入多组训练图像块,每组训练图像块包括聚焦的图像块 $P_f$ 与散焦的图像块 $P_d$ ;

步骤2.2,对于输入的各训练图像组 $Y = \{y_1, \dots, y_n\}$ ,进行向量化, $y_1, \dots, y_n$ 表示一个训练图像组中的具体图像块, $n$ 表示一个训练图像组中图像块的数量;

并训练稀疏字典 $D$ 使得输入数据能够被表示成为如下公式,

$$\min_{x_i} \|y_i - Dx_i\|_2^2 \quad s.t. \|x_i\|_0 \leq k$$

其中, $x_i$ 表示的是利用不同字典原子相融合来拟合图像 $y_i$ 时不同原子的权重参数, $k$ 表示 $x_i$ 稀疏程度的值;

步骤2.3,对于步骤2.2所得稀疏字典 $D$ ,读取步骤1.1输入的待检测的图像数据信息进行计算,获得模糊估计图 $S_D$ ;

步骤2.4,绘制步骤2.3所得模糊估计图 $S_D$ 所对应的灰度直方图 $H_D$ ,并根据灰度直方图分析得到模糊估计图的最大值 $I_{max}$ 与最小值 $I_{min}$ ;

步骤2.5,利用步骤2.4所得模糊估计图 $S_D$ 的最大值 $I_{max}$ 与最小值 $I_{min}$ 计算图像的Michelson对比度 $C$ ,计算公式如下,

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

步骤2.6,对于步骤2.5计算得到的Michelson对比度 $C$ ,利用如下公式映射到预设范围以内,

$$\lambda = \frac{1}{1 + e^{-(a+bC)}}$$

其中, $\lambda$ 为摄影师拍摄意图参数,参数 $a$ 和参数 $b$ 的取值与预设范围相应;

步骤2.7,根据步骤2.3所得模糊估计图 $S_D$ ,获取待检测图像中的聚焦区域,用聚焦区域的局部最大值点作为视觉定位点 $F_i$ ,对于每一个视觉定位点进行高斯平滑处理,得到基于视觉定位点 $F_i$ 的视觉密度图 $D_i$ ,将所有视觉定位点 $F_i$ 对应的视觉密度图 $D_i$ 叠加得到基于模糊估计融合的显著性特征图,记为高层视觉的显著性特征图 $S_T$ ;

步骤2.8,利用步骤2.6所得摄影师拍摄意图参数 $\lambda$ ,将其作为融合权重,指导不同机制下显著性特征图的融合,融合公式如下所示,

$$S = (1-\lambda)S_B + \lambda S_T$$

其中, $S$ 表示得到的图像显著性检测结果, $S_B$ 为低层视觉的显著性特征图, $S_T$ 为高层视觉的显著性特征图。

2. 根据权利要求1所述基于模糊估计融合的图像显著性检测方法,其特征在于:步骤

2.6中,预设范围为[0.2,0.8]。

3.根据权利要求2所述基于模糊估计融合的图像显著性检测方法,其特征在于:参数a的取值为-1.38,参数b的取值为2.76。

## 基于模糊估计融合的图像显著性检测方法

### 技术领域

[0001] 本发明涉及图像显著性检测领域,具体涉及一种基于模糊估计的图像显著性检测方法。

### 背景技术

[0002] 近年来,随着数字技术的不断发展,照片尺寸越来越大,分辨率越来越高,照片所蕴含的信息也变得越来越丰富,想要及时处理图像中的全部信息对于计算机图像系统而言是个不小的挑战。因此需要有一种选择机制来协助对图像中的信息进行选择性处理,以减轻图像处理系统的负担。人类的视觉显著性可以帮助人类视觉系统分辨图像中蕴含重要信息的区域,同时降低背景区域的干扰。通过将视觉显著性模型应用于计算机图像处理领域,可以更加快速的提取图像中的信息,提高图像处理精度。因此,如何模拟人类的视觉观察过程,获得更加高效精确的图像显著性检测方法,成为当前计算机视觉领域亟待解决的课题。

[0003] 目前,用于图像显著性检测的方式从原理上可以分为两种,一种是利用低层视觉特征,如颜色特征,方向特征以及对比度特征等的自底向上的显著性检测方式。这种检测方式基于图像的数据特征,具有检测速率快,计算复杂度低等多重优势。但这种传统的显著性检测方式容易产生检测错误,因而很少单独使用。另一种是利用高层视觉特征,即利用先验知识协助对图像进行显著性检测的自顶向下的显著性检测方式。这种显著性检测方式检测精度高,而且能够随着不同的检测任务进行灵活的调整,但是计算复杂度较高,耗时较长。当前显著性检测领域最为热门的研究在于将自底向上的显著性检测方式与自顶向下的显著性检测方式相融合,既利用图像的低层视觉特征也利用图像的高层视觉特征,汲取上述两种显著性检测方式的优点,获得更加高效精确的显著性检测结果。但当前流行的融合自底向上与自顶向下的显著性检测方式未涉及利用摄影师拍摄时的心理特征协助进行显著性检测,而摄影师的拍摄心理对于图像的理解具有重大的意义。因此,本发明针对当前显著性检测方法存在的不足,创造性的提出量化摄影师拍摄时的心理意图并将其作为先验知识辅助进行图像显著性检测。

### 发明内容

[0004] 本发明针对传统图像显著性检测方法存在的不足,提出一种基于模糊估计融合的图像显著性检测方法的技术方案。

[0005] 本发明技术方案提供一种基于模糊估计融合的图像显著性检测方法,包括低层视觉特征获取阶段与模糊度特征应用阶段;

[0006] 所述低层视觉特征获取阶段包含以下步骤,

[0007] 步骤1.1,输入待检测的测试图像,

[0008] 步骤1.2,根据步骤1.1输入的测试图像,利用自底向上的图像显著性检测算法获取图像的低层视觉特征,生成待检测图像对应的低层视觉的显著性特征图,用 $S_B$ 表示;

[0009] 所述模糊度特征应用阶段包含以下步骤:

[0010] 步骤2.1,输入多组训练图像块,每组训练图像块包括聚焦的图像块 $P_f$ 与散焦的图像块 $P_d$ ;

[0011] 步骤2.2,对于输入的各训练图像组 $Y = \{y_1, \dots, y_n\}$ ,进行向量化, $y_1, \dots, y_n$ 表示一个训练图像组中的具体图像块, $n$ 表示一个训练图像组中图像块的数量;

[0012] 并训练稀疏字典 $D$ 使得输入数据能够被表示成为如下公式,

$$[0013] \quad \min_{x_i} \|y_i - Dx_i\|_2^2 \quad s.t. \|x_i\|_0 \leq k$$

[0014] 其中, $x_i$ 表示的是利用不同字典原子相融合来拟合图像 $y_i$ 时不同原子的权重参数, $k$ 表示 $x_i$ 稀疏程度的值;

[0015] 步骤2.3,对于步骤2.2所得稀疏字典 $D$ ,读取步骤1.1输入的待检测的图像数据信息进行计算,获得模糊估计图 $S_D$ ;

[0016] 步骤2.4,绘制步骤2.3所得模糊估计图 $S_D$ 所对应的灰度直方图 $H_D$ ,并根据灰度直方图分析得到模糊估计图的最大值 $I_{max}$ 与最小值 $I_{min}$ ;

[0017] 步骤2.5,利用步骤2.3所得模糊估计图 $S_D$ 的最大值 $I_{max}$ 与最小值 $I_{min}$ 计算图像的Michelson对比度 $C$ ,计算公式如下,

$$[0018] \quad C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

[0019] 步骤2.6,对于步骤2.5计算得到的Michelson对比度 $C$ ,利用如下公式映射到预设范围以内,

$$[0020] \quad \lambda = \frac{1}{1 + e^{-(a+bC)}}$$

[0021] 其中, $\lambda$ 为摄影师拍摄意图参数,参数 $a$ 和参数 $b$ 的取值与预设范围相应;

[0022] 步骤2.7,根据步骤2.3所得模糊估计图 $S_D$ ,获取待检测图像中的聚焦区域,用聚焦区域的局部最大值点作为视觉定位点 $F_i$ ,对于每一个视觉定位点进行高斯平滑处理,得到基于视觉定位点 $F_i$ 的视觉密度图 $D_i$ ,将所有视觉定位点 $F_i$ 对应的视觉密度图 $D_i$ 叠加得到基于模糊估计融合的显著性特征图,记为高层视觉的显著性特征图 $S_T$ ;

[0023] 步骤2.8,利用步骤2.6所得摄影师拍摄意图参数 $\lambda$ ,将其作为融合权重,指导不同机制下显著性特征图的融合,融合公式如下所示,

$$[0024] \quad S = (1-\lambda)S_B + \lambda S_T$$

[0025] 其中, $S$ 表示得到的图像显著性检测结果, $S_B$ 为低层视觉的显著性特征图, $S_T$ 为高层视觉的显著性特征图。

[0026] 而且,步骤2.6中,预设范围为 $[0.2, 0.8]$ 。

[0027] 而且,参数 $a$ 的取值为 $-1.38$ ,参数 $b$ 的取值为 $2.76$ 。

[0028] 本发明与现有技术相比具有以下优点:

[0029] 1.本发明充分考虑摄影师拍摄时的心理活动对于图像拍摄效果的影响,将摄影心理学引入协助进行图像的显著性检测,更加符合人眼视觉系统的感知过程。

[0030] 2.本发明创新性的引入摄影师的拍摄意图参数,并将其用于协助不同显著性信息的融合,具有更好的理论基础,结果更加准确。

[0031] 3.本发明提出的显著性检测方法能够显著提升图像的显著性检测精度,并具有较

强的鲁棒性,具有一定的推广意义。

### 具体实施方式

[0032] 本发明提出的方法将自底向上的显著性检测方式与自顶向下的显著性检测方式相融合,提出了对摄影师拍摄时的心理进行模拟与量化的框架,并将模拟与量化的结果用于指导不同机制下的显著性信息的融合,综合利用图像的低层视觉特征与模糊度特征进行检测。

[0033] 本发明首先依据输入的待检测的图像数据,计算图像的低层视觉特征图;其次模拟摄影师拍摄时的心理特征进行图像的模糊估计,计算浏览者的视觉定位点,得到对应的视觉密度图作为图像的高层视觉特征图。随后利用图像模糊估计结果量化摄影师拍摄时的意图参数,强化摄影师想要表达(也即是浏览者最容易关注到)的物体,利用量化结果指导高层视觉特征图与低层视觉特征图的融合,得到最终的图像显著性检测结果。本发明模拟摄影师拍摄时的心理特征,符合图像拍摄意图与浏览者观感之间的内在关系,所得到的图像显著性检测结果更加符合人类视觉显著性检测模式。同时,综合利用了图像的低层视觉特征与高层视觉特征,具有更好的鲁棒性与更高的检测精度。

[0034] 本发明技术方案可采用计算机软件方式支持自动运行流程,以下结合实施例详细说明本发明技术方案。

[0035] 实施例包括低层视觉特征获取阶段与模糊度特征应用阶段。

[0036] 所述低层视觉特征获取阶段包含以下步骤:

[0037] 步骤1.1,输入待检测的测试图像;

[0038] 步骤1.2,根据步骤1.1输入的测试图像,利用经典的自底向上的图像显著性检测算法获取图像的低层视觉特征,生成待检测图像对应的低层视觉显著性特征图,并用 $S_B$ 表示。

[0039] 通过以上低层视觉特征获取阶段,将本实施例中待检测的测试图像全部进行处理,获取对应的低层视觉显著性特征图,用于后续不同机制下显著性特征的融合。

[0040] 经典的自底向上的图像显著性检测算法可参见以下文献,本发明不予赘述:

[0041] [1]X.Hou and L.Zhang,“Saliency detection:a spectral residual approach,”in Proc.CVPR,2007.

[0042] [2]R.Achanta,S.Hemami,F.Estrada,and S.Susstrunk,“Frequency-tuned salient region detection,”in Proc.CVPR,2009.

[0043] 所述模糊度特征应用阶段包含以下步骤:

[0044] 步骤2.1,考虑到摄影师在拍摄图像的过程中,往往通过改变焦距来强化所想要拍摄的物体,所以,越聚焦的物体越有可能是摄影师想要强化的内容。因此,本方法考虑利用图像的模糊估计的方式来评估图像中摄影师想要强化的内容。输入多组不同的训练图像块,每组训练图像块包括聚焦的图像块 $P_f$ 与散焦的图像块 $P_a$ ;实施例采用尺寸为 $8 \times 8$ 的图像块进行训练。

[0045] 所述训练图像块来源于对随机抓取的数以千计的图像进行分割,并对图像块进行一定程度的高斯模糊处理。具体实施时可由本领域技术人员自行预先选择聚焦的图像块和散焦的图像块。

[0046] 步骤2.2,对于输入的各训练图像组 $Y = \{y_1, \dots, y_n\}$ ,进行向量化, $y_1, \dots, y_n$ 表示一个训练图像组中的具体图像块, $n$ 表示一个训练图像组中图像块的数量,每个训练图像组中包括聚焦的图像块与散焦的图像块。

[0047] 并训练稀疏字典 $D$ 使得输入数据可以被表示成为如下公式。

$$[0048] \quad \min_{x_i} \|y_i - Dx_i\|_2^2 \quad s.t. \|x_i\|_0 \leq k$$

[0049] 其中, $x_i$ 表示的是利用不同字典原子相融合来拟合图像 $y_i$ 时不同原子的权重参数, $k$ 为表示 $x_i$ 稀疏程度的值。公式能够帮助拟合结果更加逼近真实图像数据,使得图像模糊估计结果更加精确。

[0050] 步骤2.3,对于获得聚焦散焦稀疏字典 $D$ ,读取步骤1.1输入的待检测的图像数据信息进行计算,获得模糊估计图 $S_D$ 。

[0051] 在模糊估计图 $S_D$ 中,每个像素点对应的灰度值用构成该像素点的字典原子数表示。构成的字典原子数越多,表示该像素点所对应的物体越清晰,构成的字典原子数越少,表示对应的物体越模糊。

[0052] 计算获得模糊估计图可参见以下文献,本发明不予赘述:

[0053] [3] J. Shi, L. Xu, and J. Jia, "Just noticeable defocus blur detection and estimation," in Proc. CVPR, 2015.

[0054] 步骤2.4,绘制模糊估计图 $S_D$ 所对应的灰度直方图 $H_D$ ,并根据灰度直方图分析得到模糊估计图的最大值 $I_{max}$ 与最小值 $I_{min}$ 。

[0055] 灰度直方图 $H_D$ 中,灰度越小表示图像内容越模糊,灰度越大表示图像内容越清晰。同一个物体往往由具有相近模糊程度的一组像素点构成,而不同的物体由于到摄像机镜头的距离差异而往往具有不同的模糊程度。因此,灰度直方图中呈现出不同的波峰与波谷。波峰表示具有相近模糊程度的像素点的集合,因而对应同一类物体,而波谷则表示不同物体之间的模糊度差异,可以用于区分不同的物体。在本方法中,最靠近灰度直方图原点的第一个波峰对应的峰值被认为代表平均背景像素模糊度,记录为 $I_{min}$ ,而离灰度直方图原点最远的波峰对应的峰值被认为代表图像中最聚焦的物体的平均像素模糊度,记录为 $I_{max}$ 。

[0056] 步骤2.5,利用模糊估计图 $S_D$ 的最大值 $I_{max}$ 与最小值 $I_{min}$ 计算图像的Michelson对比度 $C$ (能见度),计算公式如下:

$$[0057] \quad C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

[0058] 步骤2.6,对于步骤2.5计算得到的Michelson对比度 $C$ ,在本实施例中将其数值利用如下公式映射到 $[0.2, 0.8]$ 的范围以内。具体实施时,映射范围可以根据图像数据特征情况进行相应调整,本领域技术人员可自行预设,应控制在 $[0, 1]$ 的范围以内。

$$[0059] \quad \lambda = \frac{1}{1 + e^{-(a+bC)}}$$

[0060] 对于计算得到的数值 $\lambda$ ,将其定义为摄影师的拍摄意图参数,用来描述摄影师拍摄时的摄影意图在图像中的展现情况。本实施例中,为了将Michelson对比度 $C$ 值控制在 $[0.2, 0.8]$ 以内,参数 $a$ 的取值为 $-1.38$ ,参数 $b$ 的取值为 $2.76$ ,具体实施时本领域技术人员可自行预设参数 $a$ 、 $b$ 的取值。

[0061] 步骤2.7,根据步骤2.3计算获得的模糊估计图 $S_D$ ,具体实施时可以设定阈值获取待检测图像中的聚焦区域,在本实施例中,阈值取值为35,灰度值大于35的区域被认定为待检测图像的聚焦区域。用模糊估计图 $S_D$ 中聚焦区域的局部最大值点表示该区域最为聚焦的地方(即最有可能被浏览者的视觉系统关注到的视觉定位点),记录为 $F_i$ ,对于每一个视觉定位点,对其进行高斯平滑处理,得到基于该视觉定位点 $F_i$ 的视觉密度图(Density Map),记录为 $D_i$ ,将所有视觉定位点 $F_i$ 对应的视觉密度图 $D_i$ 叠加得到基于模糊估计融合的显著性特征图,记录为 $S_T$ 。

[0062] 步骤2.8,利用步骤2.6获得的摄影师拍摄意图参数 $\lambda$ ,将其作为融合权重,指导不同机制下(低层视觉特征与模糊度特征)显著性特征图的融合,融合公式如下所示:

$$[0063] \quad S = (1-\lambda) S_B + \lambda S_T$$

[0064] 其中, $S$ 表示得到的图像显著性检测结果, $S_B$ 为低层视觉显著性特征图, $S_T$ 为高层视觉显著性特征图, $\lambda$ 为摄影师的拍摄意图参数。

[0065] 本文中所描述的具体实施例仅仅是对本发明精神作举例说明,本发明所属技术领域的技术人员可对所描述的具体实施例做各种各样的修改或补充或采用类似的方式替代,但不会偏离本发明的精神或超越所附权利要求书所定义的范围。