



(12) 发明专利

(10) 授权公告号 CN 110689599 B

(45) 授权公告日 2023.05.19

(21) 申请号 201910849986.2

G06N 3/0455 (2023.01)

(22) 申请日 2019.09.10

G06N 3/0464 (2023.01)

(65) 同一申请的已公布的文献号

G06N 3/0475 (2023.01)

申请公布号 CN 110689599 A

G06N 3/084 (2023.01)

(43) 申请公布日 2020.01.14

G06N 3/094 (2023.01)

G06N 3/096 (2023.01)

(73) 专利权人 上海大学

(56) 对比文件

地址 200444 上海市宝山区上大路99号

CN 106845471 A, 2017.06.13

(72) 发明人 王永芳 王宇彤

CN 107346436 A, 2017.11.14

(74) 专利代理机构 上海上大专利事务所(普通合伙) 31205

郑一欢 等. 基于区域协方差矩阵融合深度的显著性检测方法.《工业控制计算机》.2018, 第31卷(第9期),

专利代理师 陆聪明

审查员 么旭君

(51) Int. Cl.

G06V 10/46 (2022.01)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

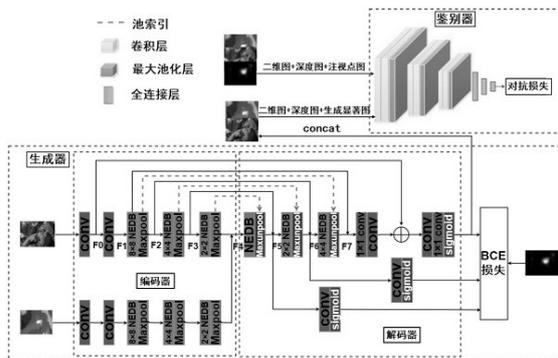
权利要求书2页 说明书7页 附图2页

(54) 发明名称

基于非局部增强的生成对抗网络的3D视觉显著性预测方法

(57) 摘要

本发明公开了一种基于非局部增强的生成对抗网络的3D视觉显著性预测方法。具体步骤为：(1)非局部密集增强的编解码器网络生成器：输入是RGB图和深度图，经过网络提取图像的多尺度特征，非局部增强网络可结合局部信息与全局信息和实现特征重用。(2)判别网络：将RGB图、深度图和显著性图作为鉴别器的输入，输出表示输入的图片为人眼显著性的概率。(3)基于非局部增强的生成对抗网络模型训练：在训练过程中，使用双路编码器分别对RGB图像和深度图像分别训练，且深度图像采用迁移训练；采用了多尺度生成损失函数和对抗损失函数联合作为最终损失函数来指导训练。(4)、最后，利用步骤(3)训练好的网络模型进行测试，可得到显著性图。



1. 一种基于非局部增强的生成对抗网络的3D视觉显著性预测方法,其特征在于,包括如下步骤:

步骤1、非局部密集增强的编解码器网络生成器:输入是RGB图和深度图,学习对象是人眼注视数据;经过非局部增强网络提取图像的多尺度特征,非局部增强网络结合局部信息与全局信息实现特征重用,通过特征融合,使RGB-D输入生成显著性图;

步骤2、判别网络:将RGB图、深度图和显著性图级联一起形成五通道图像,作为鉴别器的输入,输出表示输入的图片为人眼视图的概率,如果为1,就代表百分之百是人眼视图,而输出为0,就代表不可能是人眼视图;

步骤3、基于非局部增强的生成对抗网络模型训练:在训练过程中,使用双路编码器分别对RGB图和深度图像分别训练,其中深度图像采用迁移训练;生成网络的任务就是尽可能生成符合实际的显著性图去欺骗判别网络,而判别网络负责把生成网络生成的显著性预测图和人眼视图分别开来,采用多尺度生成损失函数和对抗损失函数联合作为最终损失函数来指导训练,训练完成后得到3D视觉显著性预测的模型;

步骤4、利用步骤3训练好的网络模型进行测试,输入RGB图和深度图,得到最终的生成显著性图;

所述步骤1中的多尺度特征提取,具体步骤如下:

第一个卷积层为:卷积核 $3 \times 3$ ,步长为1,填充为1,输出F0第一个浅特征提取,尺寸与原图像尺寸相同,深度为64;将浅特征F0送到第二个卷积层以获得浅特征F1,将浅层特征F0与靠近整个网络出口的层连接起来,实现促进梯度反向传播和像素预测的残差学习;然后使用绕过中间层的远程跳过连接,将第一个 $8 \times 8$ 非局部增强密集块输出;特征F2,特征F3都和之后的编码器网络层连接起来,这种跳过连接提供远程信息补偿,使得原始像素值和低级别特征激活在整个架构的最后仍然可用;通过最大池化层和最大上池化操作,中间特征激活的空间大小在编码阶段逐渐减小,在解码阶段逐渐增大;同时构建编码和解码层时采用了多尺度的非局部操作,这样也实现了多尺度特征的提取;

所述步骤3中的深度图像的迁移训练及多尺度生成损失函数和对抗损失函数联合作为最终损失函数,具体步骤如下:

3.1深度图像的迁移训练:由于没有足够的深度训练数据,因此无法完全调整整个网络;首先迁移训练好的RGB编码器网络的参数初始化深度编码器网络,因为两个特征提取编码器的任务是相同的,所以会达到效果;另外,网络的浅层提取边缘特征,而RGB图像和深度图像所表现的边缘特征不同,对与无结构的边缘主要依靠颜色区分,对于无纹理的边缘主要依靠结构深度区分;所以选择从浅层开始对深度特征编码器进行微调;

3.2多尺度生成损失函数和对抗损失函数联合:针对多尺度损失,从不同的解码器层中提取特征,提取编码器第一个最大上池化层的输出F5和第二个最大上池化层的输出F6,以及最后的生成器输出,形成不同大小的输出,提出的多尺度损失函数为:

$$L_M(\{S\}, \{T\}) = \sum_{i=1}^M \lambda_i L_{BCE}(S_i, T_i) \quad (1)$$

其中 $S_i$ 表示从解码器层提取的第 $i$ 个输出, $T_i$ 表示具有与 $S_i$ 相同尺度的人眼注视点图, $M$ 为尺度的个数, $\lambda$ 是不同尺度的权重,把更多的权重放在更大的尺度上,具体地说,使用F5、

F6以及生成器最终的输出,它们的尺寸分别为原来尺寸的1/4、1/2、1, $\lambda$ 设置为0.6、0.8、1;

$$L_{\text{BCE}}(S_i, T_i) = -\frac{1}{N} \sum_{i=1}^N S_i \log(T_i) + (1 - S_i) \log(1 - T_i) \quad (2)$$

其中 $L_{\text{BCE}}(S_i, T_i)$ 为内容损失函数,采用二进制交叉熵损失函数,即所有像素上的二进制交叉熵的平均值,其中N为总像素值。

2. 根据权利要求1所述的基于非局部增强的生成对抗网络的3D视觉显著性预测方法,其特征在于,所述步骤2中的判别网络,具体步骤如下:

判别网络由卷积层、池化层和最后的全连接层组成,卷积层都使用ReLU激活,而完全连接的层使用tanh激活,只有最后一层使用sigmoid激活。

## 基于非局部增强的生成对抗网络的3D视觉显著性预测方法

### 技术领域

[0001] 本发明涉及一种三维(3D)视觉显著性预测方法,特别是涉及一种基于非局部增强的生成对抗网络的3D视觉显著性预测方法,属于计算机图像处理技术领域,可用于计算机视觉领域中的图像预处理过程。

### 背景技术

[0002] 对人类视觉系统来说,它无法同时处理所有的视觉信息,而是选择性地注意到一些重要的信息,以便优先处理,这种视觉信息处理机制是人类视觉系统最重要的特征之一,被称为视觉显著性。对图像进行视觉显著性分析的过程也被称为显著性预测。目前,针对2D场景的显著区域预测,已有大量模型提出,预测结果也越来越准确,这些模型被用于各种视觉处理应用,例如重定目标、质量评估、编码等。近年来,随着立体显示技术的快速发展,3D视频技术在日常生活中越来越受欢迎,3D图像/视频的应用越来越多,如3D视频编码、3D视觉质量评估、3D目标分割等。这些新兴的3D应用增加了3D视觉显著性模型的需求。

[0003] 现有的3D视觉显著性模型大多是基于2D视觉显著性模型的。基于传统方法的3D视觉注意模型可分为三种:一种是深度加权模型,这类模型在生成3D显著性图时,利用深度信息对2D显著性预测的计算结果进行加权。2010年,Zhang等人提出一种基于多个感知刺激的用于3D视觉注意点检测的立体视觉显著性算法;一种是深度显著性模型,该方法根据深度特征与二维视觉特征分别计算显著性,然后将这些显著性图融合得到最终的三维显著性。2014年,Fang等人提出了一种考虑亮度、颜色、纹理和深度四种不同属性的立体图像显著性计算模型,他们将每个图像分割成块,并将每个块的DCT变换的DC和AC系数作为其对应的特征,生成了几个特征图,并将它们线性组合;还有一种是立体视觉模型,考虑了人类视觉系统中的立体感知机制,将立体图像的左右视图图像作为输入,不需要深度图来检测三维视觉显著性。

[0004] 由于缺乏对三维视觉感知知识的了解,这些手工提取特征的方法不能有效、准确地从原始图像中提取特征。因此,一些研究者用深度学习的方法研究3D显著性。2016年,Zhang等人采用预训练CNN模型分别生成二维和深度显著性图,然后采用线性融合方法获得了三维图像的最终显著性图。此外,中心偏置机制也被用来增强显著性映射。2018年,Liu等人在其提出的二维视觉注意模型Mr-CNN的基础上,直接将二维图像和深度图像合在一起作为输入,并将输入图像通道从三通道改为四通道。2018年,Anh-Duc Nguyen等人使用了从亮度、颜色和视差信息中提取的七个低级特征图,并将它们集成到两个基于深度学习的图像模型中,用于对失真的3D图像进行显著性预测,大大提高了显著性预测性能。

[0005] 以上这些方法虽然利用深度学习来生成3D显著性图,但对深度图和彩色中的特征提取不够,并且没有较好的将深度信息和彩色信息进行融合。

### 发明内容

[0006] 本发明的目的在于针对上述现有技术存在的不足,提出了一种基于非局部增强的

生成对抗网络的3D视觉显著性预测方法,旨在提高3D图像显著性预测的准确率。本发明方法能有效地提高3D图像显著性的客观指标,而且在主观视觉上也有更好的效果。此外,本发明对于卷积神经网络在3D视觉显著性的应用也具有重要的借鉴意义。

[0007] 为达到上述目的,本发明的采用的技术方案如下:

[0008] 一种基于非局部增强的生成对抗网络的3D视觉显著性预测方法,包括如下步骤:

[0009] 步骤1、非局部密集增强的编解码器网络生成器:输入是RGB图和深度图,学习对象是人眼注视数据;经过非局部增强网络提取图像的多尺度特征,非局部增强网络结合局部信息与全局信息和实现特征重用,通过特征融合,使RGB-D输入生成显著性图;

[0010] 步骤2、判别网络:将RGB图、深度图和显著性图级联一起形成五通道图像,作为鉴别器的输入,输出表示输入的图片为人眼注视图的概率,如果为1,就代表百分之百是人眼注视图,而输出为0,就代表不可能是人眼注视图;

[0011] 步骤3、基于非局部增强的生成对抗网络模型训练:在训练过程中,使用双路编码器分别对彩色图像和深度图像分别训练,其中深度图像的采用迁移训练;生成网络的任务就是尽可能生成符合实际的显著性图去欺骗判别网络,而判别网络负责把生成网络生成的显著性预测图和人眼注视图分别开来,采用多尺度生成损失函数和对抗损失函数联合作为最终损失函数来指导训练,训练完成后得到3D视觉显著性预测的模型;

[0012] 步骤4、利用步骤3训练好的网络模型进行测试,输入RGB图和深度图,得到最终的生成显著性图。

[0013] 本发明方法主要是使用非局部增强的生成对抗网络作为生成器,使用双路编码器分别提取彩色图像和深度的局部信息与全局信息特征,且采用密集连接实现特征重用,同时提出了多尺度损失函数,对不同尺度的显著性预测图计算损失,利用局部损失和全局损失,提高显著性预测精度。该方法分为非局部增强的编解码器网络生成器、判别网络、基于非局部增强的生成对抗网络模型训练三个部分。在非局部增强的编解码器网络生成器部分,使用局部连接块有助于结合局部信息与全局信息获得多尺度的特征,密集连接实现了特征重用,也有助于网络的梯度反向传播和收敛。在判别网络部分,使用判别网络将生成网络生成的显著性预测图和人眼注视图分别开来。生成器和判别器构成了一个动态的“博弈过程”。从而得到一个生成式的模型,它用来生成显著性图。

[0014] 本发明与现有技术相比较,具有如下显而易见的突出实质性特点和显著优点:

[0015] 1、本发明方法提出非局部增强的生成对抗网络的3D视觉显著性预测方法,该方法提出由非局部连接块和密集连接组成的非局部增强编解码器网络作为生成器,非局部连接块有助于结合局部信息与全局信息获得多尺度的特征,密集连接实现了特征重用,也有助于网络的梯度反向传播和收敛。相比于127MB的VGG16生成器参数文件,该网络使生成器模型参数文件降到4MB,实现了网络轻量化。

[0016] 2、本发明方法提出了多尺度损失函数,对不同尺度的显著性预测图计算损失,同时利用局部损失和全局损失,提高显著性预测精度

[0017] 3、本发明方法使用双路编码器分别提取彩色图像特征和深度特征,并用彩色特征提取网络的权重来初始化深度特征提取网络参数的方法来解决三维视觉显著数据集小的问题,提高了三维显著性预测的精度。

## 附图说明

[0018] 图1为本发明非局部增强的生成对抗网络的3D视觉显著性预测方法的网络结构框图。

[0019] 图2为基于非局部增强的生成器框架图。

[0020] 图3各模型预测3D图像的显著性预测主观结果图对比,第一行为RGB图,第二行为对应深度图,第三行为相应的人眼注视点图(ground truth),第四到第十行分别为SalGAN、NedM1GAN(提出二维视觉模型无深度信息)、提出三维视觉模型、GBVS模型、Itti模型、Fang模型以及Salcovar模型。

## 具体实施方式

[0021] 本发明的优选实施例结合附图详述如下:

[0022] 本实施例的非局部增强的生成对抗网络的3D视觉显著性预测网络结构如图1所示。在Ubuntu 16.04,PyTorch环境下编程仿真实现本方法。

[0023] 本方法具体包括如下步骤:

[0024] 步骤1、非局部密集增强的编解码器网络生成器:输入是RGB图和深度图,学习对象是人眼注视数据。经过非局部增强网络提取图像的多尺度特征,非局部密集增强网络助于结合局部信息与全局信息和实现特征重用,通过特征融合,使RGB-D输入生成显著性图。

[0025] 使用一个非局部密集增强的编解码器网络作为生成器,非局部连接块有助于结合局部信息与全局信息,密集连接实现了特征重用,有助于网络的梯度反向传播和收敛,如图2所示。

[0026] 使用两个卷积层来提取输入二维图像的浅特征。第一个卷积层为:卷积核 $3 \times 3$ ,步长为1,填充为1,输出F0第一个浅特征提取,尺寸与原图像尺寸相同,深度为64。接下来,将浅特征F0送到第二个卷积层以获得浅特征F1。如图2所示,将浅层特征F0与靠近整个网络出口的层连接起来,这样它们实现了促进梯度反向传播和像素预测的残差学习。然后使用绕过中间层的远程跳过连接,将第一个 $8 \times 8$ 非局部增强密集块(Non-local Enhanced Dense Block, NEDB)输出,特征F2,特征F3都和之后的编码器网络层连接起来,这种跳过连接可以提供远程信息补偿,使得原始像素值和低级别特征激活在整个架构的最后仍然可用。通过最大池化层和最大上池化操作,中间特征激活的空间大小在编码阶段逐渐减小,在解码阶段逐渐增大。因此,由于NEDB中的非局部操作需要计算特征激活图每两个空间位置之间的成对关系,当空间维数变大时,计算量急剧增加。为了解决这一问题,并在不同空间分辨率的特征激活之间构建更灵活的非局部增强,所以在构建编码和解码层时采用了多尺度的非局部操作。具体来说,对于空间分辨率最低的特征激活(如图2中的F4),后续的NEDB直接作用于整个特征激活图,实现全局非局部增强。对于空间分辨率较高的特征激活,首先将其划分为区域网格(如图2所示, $k \times k$  NEDB表示在执行区域非局部操作之前如何划分输入特征图)。例如F1被划分为一个 $8 \times 8$ 的网格,然后让后续的NEDB在每个区域内进行特征激活。因此,这种区域级的非局部增强能够防止由于直接处理高分辨率的特征激活而导致不可接受的计算消耗。另一方面,与传统的局部卷积运算相比,区域级非局部增强能够检索到长期的结构线索。

[0027] 步骤2、判别网络:将RGB图、深度图和显著性图级联一起形成五通道图像,作为鉴

别器的输入,输出表示输入的图像为人眼注视图的概率,如果为1,就代表百分之百是人眼注视图,而输出为0,就代表不可能是人眼注视图。

[0028] 判别器能准确的把生成的图片和真实的图片分类,将图像和显著性图输入判别器,当使用生成的对抗性网络生成符合实际的图像时,则输出1,否则对生成的图片输出0。在对生成函数参数进行更新时,损失函数结合了判别器的误差和交叉熵对人眼注视点图的影响,提高了反向训练的稳定性和收敛速度。本发明的判别器网络由卷积层、池化层和最后的全连接层组成,卷积层都使用ReLU激活,而完全连接的层使用tanh激活,只有最后一层使用sigmoid激活,其网络框架如表1所示。

[0029] 表1鉴别器网络框架设计

层	深度	卷积核	步长	pad	激活函数
Conv1_1	3	1×1	1	1	RELU
Conv2_2	32	3×3	1	1	RELU
Pool1	-	2×2	2	0	-
Conv2_1	64	3×3	1	1	RELU
Conv2_2	64	3×3	1	1	RELU
Pool2	-	2×2	2	0	-
Conv3_1	64	3×3	1	1	RELU
Conv3_2	64	3×3	1	1	RELU
Pool3	-	2×2	2	0	-
fc4	100	-	-	-	tanh
fc5	2	-	-	-	tanh
fc6	1	-	-	-	sigmoid

[0032] 步骤3、基于非局部增强的生成对抗网络模型训练:在训练过程中,使用双路编码器分别对彩色图像和深度图像分别训练,其中深度图像的采用迁移训练;生成网络的任务就是尽可能生成符合实际的显著性图去欺骗判别网络,而判别网络负责把生成网络生成的显著性预测图和人眼注视图分别开来,采用多尺度生成损失函数和对抗损失函数联合作为最终损失函数来指导训练,训练完成后得到3D视觉显著性预测的模型。具体步骤如下:

[0033] 3.1深度图像的迁移训练:由于没有足够的深度训练数据,因此无法完全调整整个网络;首先迁移训练好的RGB编码器网络的参数初始化深度编码器网络,因为两个特征提取编码器的任务是相似的,所以会达到比较好的效果;另外,网络的浅层主要提取边缘特征,而RGB图像和深度图像所表现的边缘特征有很大不同,比如对与无结构的边缘主要依靠颜色区分,对于无纹理的边缘主要依靠结构深度区分;所以选择从浅层开始对深度特征编码器进行微调。

[0034] 3.2多尺度生成损失函数和对抗损失函数联合:针对多尺度损失,从不同的解码器层中提取特征,如图2所示,提取编码器第一个最大上池化层的输出F5和第二个最大上池化层的输出F6,以及最后的生成器输出,形成不同大小的输出。提出多尺度损失函数为:

$$[0035] \quad L_M(\{S\}, \{T\}) = \sum_{i=1}^M \lambda_i L_{BCE}(S_i, T_i) \quad (1)$$

[0036] 其中 $S_i$ 表示从解码器层提取的第 $i$ 个输出, $T_i$ 表示具有与 $S_i$ 相同尺度的人眼注视点图(ground truth), $M$ 为尺度的个数, $\lambda$ 是不同尺度的权重,把更多的权重放在更大的尺度上,具体地说,使用F5、F6以及生成器最终的输出,它们的尺寸分别为原来尺寸的1/4、1/2、1。 $\lambda$ 设置为0.6、0.8、1。其中 $L_{BCE}(S_i, T_i)$ 为内容损失函数,采用二进制交叉熵损失函数,即所有像素上的二进制交叉熵的平均值,其中 $N$ 为总像素值,其计算公式为:

$$[0037] \quad L_{BCE}(S_i, T_i) = -\frac{1}{N} \sum_{i=1}^N S_i \log(T_i) + (1 - S_i) \log(1 - T_i) \quad (2)$$

[0038] 步骤4、利用步骤3训练好的网络模型进行测试,输入RGB图和深度图,就得到最终的生成显著性图。

[0039] 为了验证所提出的三维视觉注意模型的性能,在NUS3D-Saliency[1]、NCTU-3DFixation[2]和3D Gaze[3]三个数据集上对模型性能进行评估。本实验的环境是Ubuntu 16.04操作系统下的PyTorch平台,内存为16GB,GPU为GeForce 1080。采用了AUC-J,AUC-B,线性相关系数(CC)、Kullback-Leibler散度(KL)和相似度评分(SIM)五个显著性评价指标来评测性能。图3为各模型预测3D图像的显著性预测主观结果图对比,第一行为RGB图,第二行为其对应深度图,第三行为相应的人眼注视点图(ground truth),第四到第十行分别为SalGAN、NedMIGAN(提出二维视觉模型无深度信息)、提出三维视觉模型、GBVS模型、Itti模型、Fang模型以及Salcovar模型。

[0040] 对三维图像进行显著性预测的结果,CTU-3DFIXED[1]的测试结果如表2所示,NUS3D-Saliency[2]和3D Gaze[3]的测试结果分别如表3,表4所示,由于这两个数据集只提供了高斯模糊过的三维人眼注视点图数据,没有提供相应的点图,所以只采用了CC、KL和SIM三个显著性评价指标。其中SalGAN为[4]中的模型,NedMIGAN是本文在[4]基础上改进的二维视觉注意模型,GBVS是文献[5]中提出的二维视觉注意模型,Itti是文献[6]中的对比二维视觉注意模型,Fang是文献[7]中提出的三维视觉注意模型,Salcovar是文献[8]中提出的三维视觉注意模型。

[0041] 表2:各模型在CTU-3DFIXED数据集上的测试结果

	AUC-J	AUC-B	CC	KL	NSS	SIM
参考值	0.92	0.88	1	0	3.29	1
SalGAN	0.8273	0.7738	<b>0.5816</b>	1.3035	1.3332	<b>0.5247</b>
NedMIGAN	<b>0.8315</b>	<b>0.7892</b>	0.5741	<b>1.0137</b>	<b>1.3376</b>	0.5141
GBVS	0.7991	0.7723	0.4603	1.0394	1.1561	0.4485
Itti	0.7991	0.7702	0.4585	1.0067	1.1470	0.4511
Fang	0.7603	0.7409	0.3773	1.2270	0.9494	0.3903
Salcovar	0.7999	0.7194	0.4822	1.6020	1.1644	0.4753
本发明模型	<b>0.8429</b>	<b>0.7714</b>	<b>0.6244</b>	<b>0.8817</b>	<b>1.4125</b>	<b>0.5537</b>

[0042] 表3:各模型在NUS3D-Saliency数据集上的测试结果

	CC	KL	SIM
参考值	1	0	1
SalGAN	<b>0.4258</b>	1.9166	<b>0.2970</b>
NedMIGAN	0.4245	<b>1.8053</b>	0.2962
GBVS	0.3036	2.1865	0.1907
itti	0.3053	2.1459	0.1960
Fang	0.2711	2.3264	0.1711
Salcovar	0.3664	2.0043	0.2754
本发明模型	<b>0.4577</b>	<b>1.7132</b>	<b>0.3063</b>

[0045] 表4:各模型在3D Gaze数据集上的测试结果

	cc	kl	sim
参考值	1	0	1
SalGAN	0.5080	1.4196	0.5032
NedMIGAN	0.4665	1.4781	0.4711
GBVS	0.4333	1.6699	0.4756
itti	0.3652	1.7793	0.4678
Fang	0.3605	1.8555	0.4689
Salcovar	<b>0.5144</b>	<b>1.444</b>	<b>0.4978</b>
本发明模型	<b>0.4895</b>	<b>1.3926</b>	<b>0.4976</b>

[0047] 其中,实验结果最好的两个算法用字体加粗表示。由上述实验可见,本发明方法在3D视觉注意模型上确实有较好的鲁棒性和精确性,并且计算复杂度低,能更好地适用于实时视频质量监控。

[0048] 参考文献:

[0049] [1]Chih-Yao M,Hsueh-Ming H.Learning-based saliency model with depth information[J].Journal of Vision,2015,15(6):19.

[0050] [2]Lang C,Nguyen T V,Katti H,et al.Depth matters:influence of depth cues on visual saliency[C].European Conference on Computer Vision,2012,101-115

[0051] [3]Wang J,Dasilva M P,Lecallet P,et al.Computational Model of Stereoscopic 3D Visual Saliency[J].IEEE Transactions on Image Processing, 2013,22(6):2151-2165.

[0052] [4]Pan J,Ferrer C C,Mcguinness K,et al.SalGAN:Visual Saliency Prediction with Generative Adversarial Networks[J].arXiv.2017.

[0053] [5]Harel J,Koch C,Perona P.Graph-Based Visual Saliency[C],Advances in Neural Information Processing Systems 19,Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems,Vancouver,British

Columbia,Canada,December 4-7,2006.MIT Press,2006.

[0054] [6]Itti L,Koch C,Niebur E.A model of saliency-based visual attention for rapid scene analysis.IEEE Trans.Pattern Anal.Mach.Intell,1998,20(11): 1254-1259

[0055] [7]Fang Y,Lin W,Fang Z,et al.Learning visual saliency for stereoscopic images[C].2014IEEE International Conference on Multimedia and Expo Workshops(ICMEW).IEEE Computer Society,2014.

[0056] [8]Erdem E,Erdem A.Visual saliency estimation by nonlinearly integrating features using region covariances[J].Journal of Vision,2013,13(4):11-11.

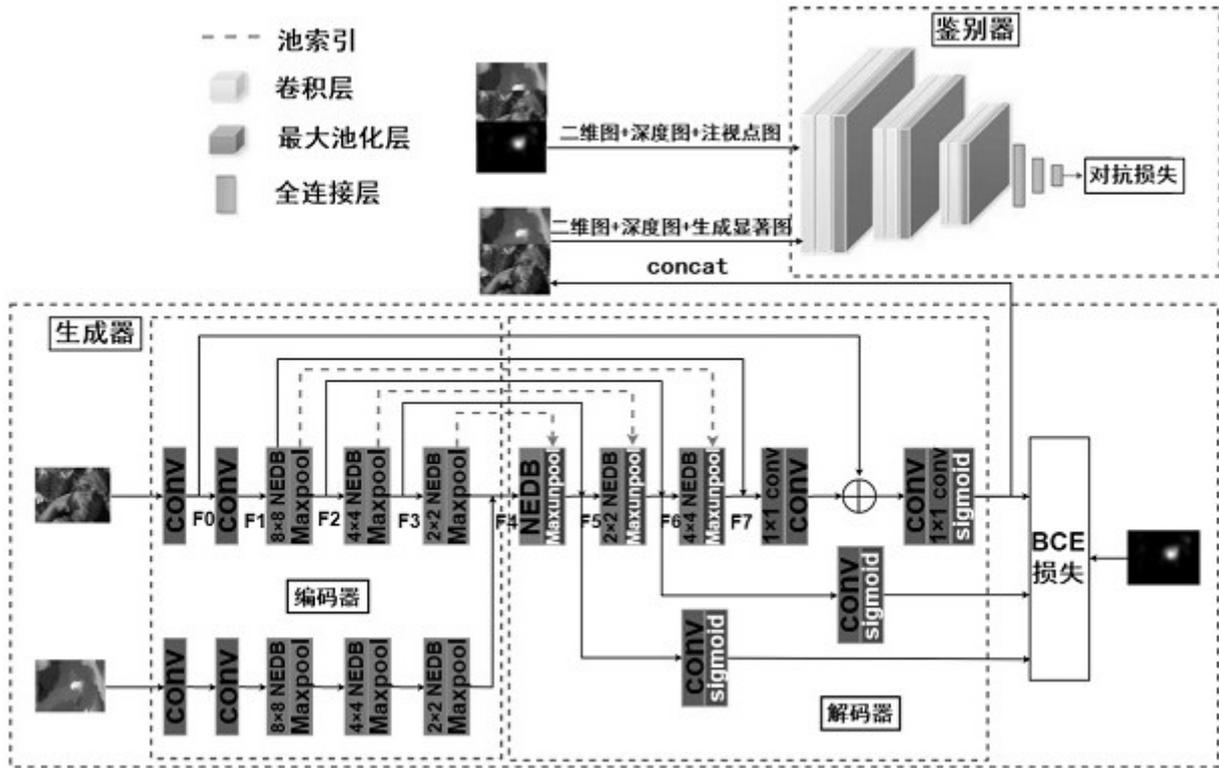


图1

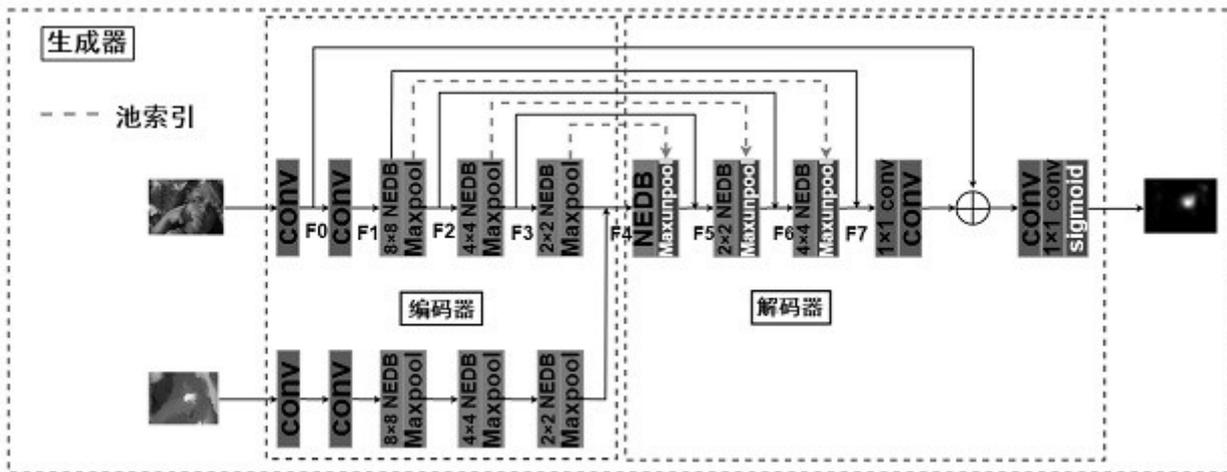


图2

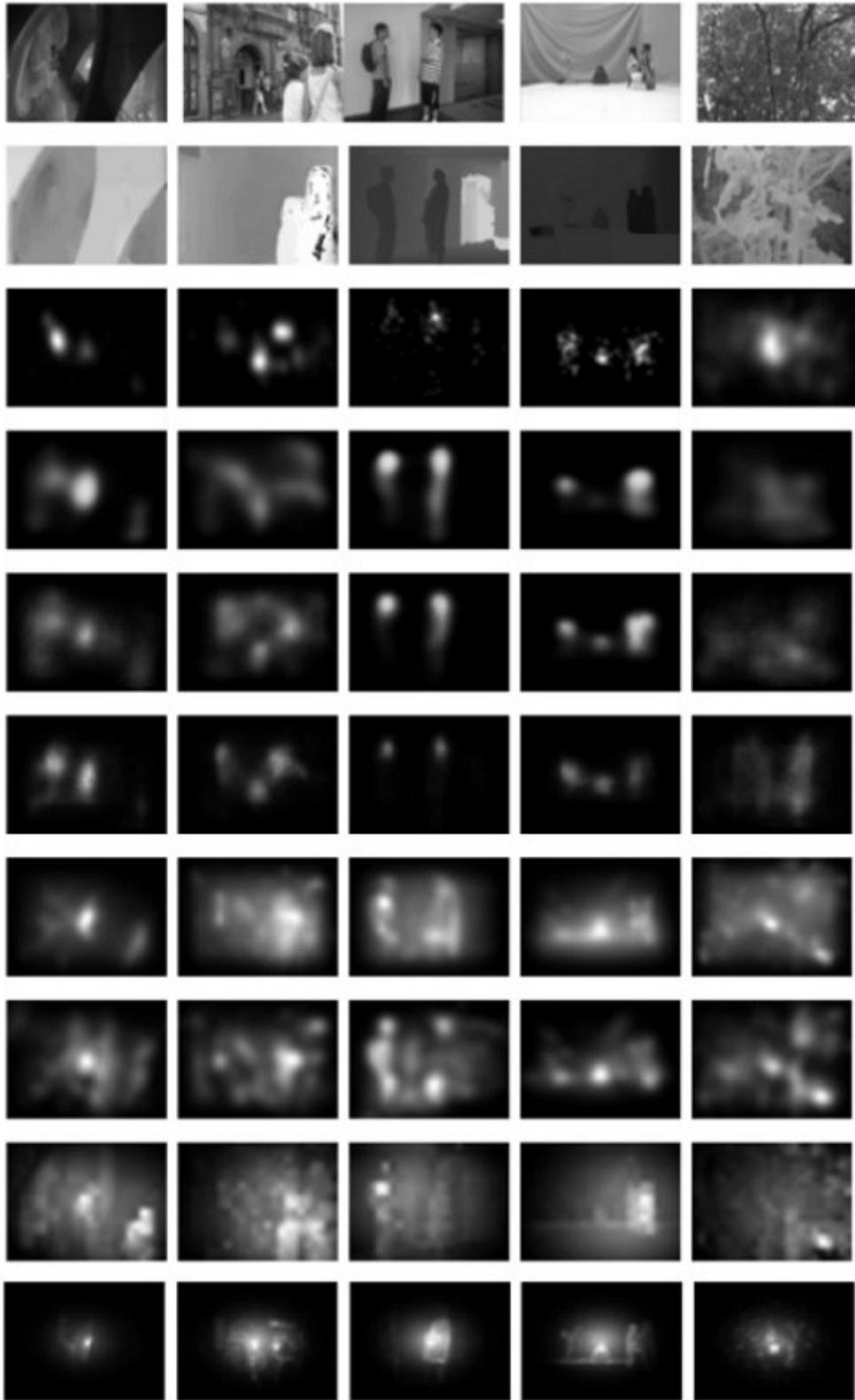


图3