



(12)发明专利申请

(10)申请公布号 CN 111832592 A
(43)申请公布日 2020.10.27

(21)申请号 201910328103.3

(22)申请日 2019.04.20

(71)申请人 南开大学

地址 300000 天津市南开区卫津路94号

申请人 华为技术有限公司

(72)发明人 程明明 曹洋 吴日辉 白蔚

(74)专利代理机构 广州三环专利商标代理有限公司 44202

代理人 郝传鑫 熊永强

(51) Int. Cl.

G06K 9/62(2006.01)

G06K 9/46(2006.01)

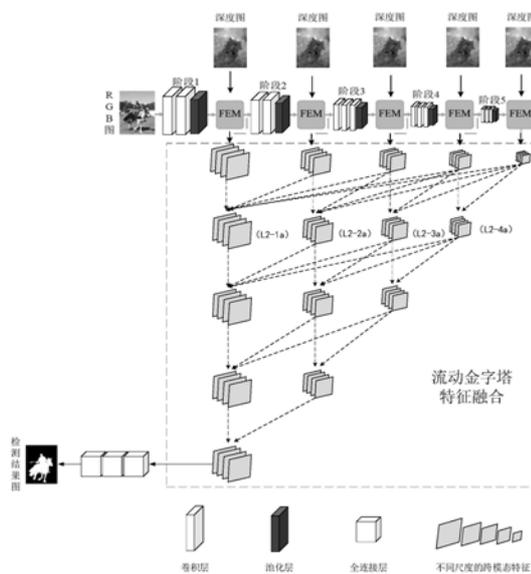
权利要求书5页 说明书25页 附图13页

(54)发明名称

RGBD显著性检测方法以及相关装置

(57)摘要

本申请提供了RGBD显著性检测方法及装置,方法包括:将RGB图像和深度图像导入到含多个特征提取模块和多个特征强化模块的神经网络模型处理,获得多个跨模态特征图;其中任意的特征提取模块用于对输入数据进行特征提取以获得该特征提取模块的处理结果;其中任意的特征强化模块用于对深度图像进行特征提取获得深度对比度增强图,并利用其优化特征提取模块的处理结果,以获得该特征强化模块的跨模态特征图。最后,利用金字塔模型对多个跨模态特征图进行融合获得显著性检测的检测结果图。实施本申请有利于在低对比度等复杂场景下取得较好的显著性检测效果。



1. 一种RGBD显著性检测方法,其特征在于,所述方法包括:

获取RGB图像和与所述RGB图像配准的深度图像;

将所述RGB图像导入到包含N个特征提取模块和N个特征强化模块的神经网络模型进行处理,以及将所述深度图像导入所述神经网络模型进行处理,获得所述N个特征强化模块的N个跨模态特征图;N为大于等于2的整数;

其中,所述N个特征提取模块中的第j个特征提取模块的处理包括:对所述第j个特征提取模块的输入数据进行特征提取,获得所述第j个特征提取模块的处理结果并输出到所述N个特征强化模块中的第j个特征强化模块;其中,当j等于1时,所述输入数据为所述RGB图像,当j不等于1时,所述输入数据为第j-1个特征强化模块输出的跨模态特征图;j为大于等于1且小于等于N的整数;

其中,所述N个特征强化模块中的第j个特征强化模块的处理包括:对导入至所述第j个特征强化模块的深度图像进行特征提取,获得深度对比度增强图;利用所述深度对比度增强图优化所述第j个特征提取模块的处理结果,获得所述第j个特征强化模块的跨模态特征图;

对所述N个跨模态特征图进行融合,获得检测结果图。

2. 根据权利要求1所述的方法,其特征在于,所述对所述第j个特征提取模块的输入数据进行特征提取,包括:

对所述第j个特征提取模块的输入数据通过至少一个卷积层处理和至少一个池化层处理。

3. 根据权利要求1或2所述的方法,其特征在于,所述对导入至所述第j个特征强化模块的深度图像进行特征提取,获得深度对比度增强图,包括:

对导入至所述第j个特征强化模块的深度图像通过卷积层集合处理以及通过激活函数处理,获得深度对比度增强图。

4. 根据权利要求1-3任一项所述的方法,其特征在于,所述利用所述深度对比度增强图优化所述第j个特征提取模块的处理结果,获得所述第j个特征强化模块的跨模态特征图,包括:

将所述深度对比度增强图与所述第j个特征提取模块的处理结果相乘,获得相乘后的结果;

将所述相乘后的结果与所述第j个特征提取模块的处理结果相加,获得所述第j个特征强化模块的跨模态特征图。

5. 根据权利要求1-4任一项所述的方法,其特征在于,所述神经网络模型还包括全连接层;所述对所述N个跨模态特征图进行融合,获得检测结果图,包括:

通过金字塔模型对所述N个跨模态特征图进行融合,获得融合结果;

将所述融合结果输出到所述全连接层进行处理,获得所述显著性检测的检测结果图。

6. 根据权利要求5所述的方法,其特征在于,所述金字塔模型包括具有多个层的金字塔,从所述金字塔的底层到所述金字塔的顶层的各个层的图像数量依次递减,所述底层的图像集包括N个图像,所述N个图像与所述N个特征提取模块输出的N个跨模态特征图一一对应;

所述通过金字塔模型对所述N个跨模态特征图进行融合,获得融合结果,包括:

按照从所述金字塔的底层到所述金字塔的顶层的顺序,获得除所述底层外的每一层的图像集;其中,除所述底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度不大于所述任一图像的尺度的所有图像融合而成,所述顶层的图像集包括一个图像;

将所述顶层的所述一个图像作为所述融合结果。

7. 根据权利要求5所述的方法,其特征在于,所述金字塔模型包括具有多个层的金字塔,从所述金字塔的底层到所述金字塔的顶层的各个层的图像数量依次递减,所述底层的图像集包括N个图像,所述N个图像与所述N个特征提取模块输出的N个跨模态特征图一一对应;

所述通过金字塔模型对所述N个跨模态特征图进行融合,获得融合结果,包括:

按照从所述金字塔的底层到所述金字塔的顶层的顺序,通过融合填充获得除所述底层外的每一层的图像集;其中,除所述底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度同所述任一图像相同的一个图像和尺度比所述任一图像小的一个图像两者融合而成,所述顶层的图像集包括一个图像;

将所述顶层的所述一个图像作为所述融合结果。

8. 根据权利要求3所述的方法,其特征在于,

所述卷积层集合中各个卷积层的卷积核参数是利用对比度增强损失函数对导入至所述第j个特征强化模块的深度图像进行监督而获得的。

9. 根据权利要求8所述的方法,其特征在于,所述利用对比度增强损失函数对导入至所述第j个特征强化模块的深度图像进行监督,包括:

将导入至所述第j个特征强化模块的深度图像通过所述卷积层集合的各个卷积层进行卷积处理,获得深度特征图;

利用对比度增强损失函数对所述深度特征图进行运算,得到所述对比度增强损失函数的值;

在所述对比度增强损失函数的值未达到局部最小值的情况下,更新所述卷积层集合的各个卷积层的卷积核参数,继续梯度下降迭代,直到达到局部最小值。

10. 根据权利要求8或9所述的方法,其特征在于,所述对比度增强损失函数 l_c 的数学描述如下所示:

$$l_c = \alpha_1 l_f + \alpha_2 l_b + \alpha_3 l_w$$

其中,

$$l_f = -\log\left(1 - 4 \times \sum_{p \in F} \frac{(p - \widehat{p}_f)^2}{N_f}\right)$$

$$l_b = -\log\left(1 - 4 \times \sum_{p \in B} \frac{(p - \widehat{p}_b)^2}{N_b}\right)$$

$$l_w = -\log(\widehat{p}_f - \widehat{p}_b)^2$$

$$\widehat{p}_f = \sum_{p \in F} \frac{p}{N_f}$$

$$\widehat{p}_b = \sum_{p \in B} \frac{p}{N_b}$$

其中, α_1 、 α_2 和 α_3 分别设置为5、5和1;

l_f 表示监督深度图像的前景部分的训练; F 为前景点集合, \hat{p}_f 为前景点均值, N_f 为前景点个数;

l_b 表示监督深度图像的背景部分的训练; B 为背景点集合, \hat{p}_b 为背景点均值, N_b 为背景点个数;

l_w 表示监督深度图像的全图训练; \hat{p}_f 为前景点均值, \hat{p}_b 为背景点均值。

11.根据权利要求1-10任一项所述的方法,其特征在于, N 等于5,5个特征提取模块具体描述如下:

第1个特征提取模块包括2个卷积层和1个池化层;

第2个特征提取模块包括2个卷积层和1个池化层;

第3个特征提取模块包括3个卷积层和1个池化层;

第4个特征提取模块包括3个卷积层和1个池化层;

第5个特征提取模块包括3个卷积层和1个池化层。

12.一种用于RGBD显著性检测的神经网络系统,其特征在于,所述神经网络系统包括神经网络模型和融合模块,所述神经网络模型包括 N 个特征提取模块和 N 个特征强化模块,所述特征提取模块和所述特征强化模块交替连接;

所述神经网络模型用于,对获取的RGB图像进行处理,以及对与所述RGB图像配准的深度图像进行处理,获得所述 N 个特征强化模块的 N 个跨模态特征图; N 为大于等于2的整数;

其中,所述 N 个特征提取模块中的第 j 个特征提取模块具体用于:对所述第 j 个特征提取模块的输入数据进行特征提取,获得所述第 j 个特征提取模块的处理结果并输出到所述 N 个特征强化模块中的第 j 个特征强化模块;其中,当 j 等于1时,所述输入数据为所述RGB图像,当 j 不等于1时,所述输入数据为第 $j-1$ 个特征强化模块输出的跨模态特征图; j 为大于等于1且小于等于 N 的整数;

其中,所述 N 个特征强化模块中的第 j 个特征强化模块具体用于:对导入至所述第 j 个特征强化模块的深度图像进行特征提取,获得深度对比度增强图;利用所述深度对比度增强图优化所述第 j 个特征提取模块的处理结果,获得所述第 j 个特征强化模块的跨模态特征图;

所述融合模块用于,对所述 N 个跨模态特征图进行融合,获得显著性检测的检测结果图。

13.根据权利要求12所述的系统,其特征在于,所述第 j 个特征提取模块具体用于:

对所述第 j 个特征提取模块的输入数据通过至少一个卷积层处理和至少一个池化层处理。

14.根据权利要求12或13所述的方法,其特征在于,所述第 j 个特征强化模块具体用于:

对导入至所述第 j 个特征强化模块的深度图像通过卷积层集合处理以及通过激活函数处理,获得深度对比度增强图。

15.根据权利要求12-14任一项所述的系统,其特征在于,所述第 j 个特征强化模块具体用于:

将所述深度对比度增强图与所述第 j 个特征提取模块的处理结果相乘,获得相乘后的结果;

将所述相乘后的结果与所述第j个特征提取模块的处理结果相加,获得所述第j个特征强化模块的跨模态特征图。

16. 根据权利要求12-15任一项所述的系统,其特征在于,所述神经网络模型还包括全连接层;

所述融合模块具体用于:

通过金字塔模型对所述N个跨模态特征图进行融合,获得融合结果;

将所述融合结果输出到所述全连接层进行处理,获得所述显著性检测的检测结果图。

17. 根据权利要求16所述的系统,其特征在于,所述金字塔模型包括具有多个层的金字塔,从所述金字塔的底层到所述金字塔的顶层的各个层的图像数量依次递减,所述底层的图像集包括N个图像,所述N个图像与所述N个特征提取模块输出的N个跨模态特征图一一对应;

所述融合模块具体用于:

按照从所述金字塔的底层到所述金字塔的顶层的顺序,获得除所述底层外的每一层的图像集;其中,除所述底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度不大于所述任一图像的尺度的所有图像融合而成,所述顶层的图像集包括一个图像;

将所述顶层的所述一个图像作为所述融合结果。

18. 根据权利要求16所述的系统,其特征在于,所述金字塔模型包括具有多个层的金字塔,从所述金字塔的底层到所述金字塔的顶层的各个层的图像数量依次递减,所述底层的图像集包括N个图像,所述N个图像与所述N个特征提取模块输出的N个跨模态特征图一一对应;

所述融合模块具体用于:

按照从所述金字塔的底层到所述金字塔的顶层的顺序,通过融合填充获得除所述底层外的每一层的图像集;其中,除所述底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度同所述任一图像相同的一个图像和尺度比所述任一图像小的一个图像两者融合而成,所述顶层的图像集包括一个图像;

将所述顶层的所述一个图像作为所述融合结果。

19. 根据权利要求14所述的系统,其特征在于,

所述卷积层集合中各个卷积层的卷积核参数是所述第j个特征强化模块利用对比度增强损失函数对导入至所述第j个特征强化模块的深度图像进行监督而获得的。

20. 根据权利要求19所述的系统,其特征在于,所述第j个特征强化模块具体用于:

将导入至所述第j个特征强化模块的深度图像通过所述卷积层集合的各个卷积层进行卷积处理,获得深度特征图;

利用对比度增强损失函数对所述深度特征图进行运算,得到所述对比度增强损失函数的值;

在所述对比度增强损失函数的值未达到局部最小值的情况下,更新所述卷积层集合的各个卷积层的卷积核参数,继续梯度下降迭代,直到达到局部最小值。

21. 根据权利要求19或20所述的系统,其特征在于,所述对比度增强损失函数 l_c 的数学描述如下所示:

$$l_c = \alpha_1 l_f + \alpha_2 l_b + \alpha_3 l_w$$

其中,

$$l_f = -\log(1 - 4 \times \sum_{p \in F} \frac{(p - \hat{p}_f)^2}{N_f})$$

$$l_b = -\log(1 - 4 \times \sum_{p \in B} \frac{(p - \hat{p}_b)^2}{N_b})$$

$$l_w = -\log(\hat{p}_f - \hat{p}_b)^2$$

$$\hat{p}_f = \sum_{p \in F} \frac{p}{N_f}$$

$$\hat{p}_b = \sum_{p \in B} \frac{p}{N_b}$$

其中, α_1 、 α_2 和 α_3 分别设置为5、5和1;

l_f 表示监督深度图像的前景部分的训练; F 为前景点集合, \hat{p}_f 为前景点均值, N_f 为前景点个数;

l_b 表示监督深度图像的背景部分的训练; B 为背景点集合, \hat{p}_b 为背景点均值, N_b 为背景点个数;

l_w 表示监督深度图像的全图训练; \hat{p}_f 为前景点均值, \hat{p}_b 为背景点均值。

22. 根据权利要求12-21任一项所述的系统,其特征在于, N 等于5,5个特征提取模块具体描述如下:

- 第1个特征提取模块包括2个卷积层和1个池化层;
- 第2个特征提取模块包括2个卷积层和1个池化层;
- 第3个特征提取模块包括3个卷积层和1个池化层;
- 第4个特征提取模块包括3个卷积层和1个池化层;
- 第5个特征提取模块包括3个卷积层和1个池化层。

23. 一种计算节点,其特征在于,所述计算节点包括处理器以及耦合的存储器;其中:所述存储器用于存储RGB图像和深度图像,还用于存储程序指令;所述处理器用于调用所述程序指令,以实现如权利要求1-11任一项所描述的方法。

24. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行如权利要求1-11任一项所描述的方法。

RGBD显著性检测方法以及相关装置

技术领域

[0001] 本申请涉及人工智能领域,尤其涉及RGBD显著性检测方法以及相关装置。

背景技术

[0002] 人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说,人工智能是计算机科学的一个分支,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式作出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。人工智能领域的研究包括机器人,自然语言处理,计算机视觉,决策与推理,人机交互,推荐与搜索, AI基础理论等。

[0003] 人类的视觉感知系统在对场景进行感知时,除了获取颜色、形状等外貌信息外,还可以感知场景的深度信息,即景深。随着成像设备的发展,基于此原理开发的RGBD (Red Green Blue Depth, 红色绿色蓝色深度)摄像头可用来获取RGB图像信息和深度信息。RGB图像信息可提供拍摄场景的颜色、光照等信息,深度信息可提供该拍摄场景的景深、立体形状、位置距离等信息。目前针对RGBD数据已开展了广泛的研究,例如RGBD显著性检测。

[0004] 作为计算机视觉领域的一项重要预处理技术,视觉显著性检测已广泛应用于检测、编码、增强、压缩等领域。在一张图片中,检测最具备视觉显著性的物体或区域是显著性检测的主要研究目标。它基于生物学理论中的人眼注意机制,从RGB图像中提取特征,对特征处理后得到显著性检测结果。

[0005] 随着深度学习的发展,卷积神经网络被应用于显著性检测领域,通过将网络模型在大体量数据集上训练提高了显著性检测的效果。然而,现有显著性检测方案中,比如低对比度、小物体、多个物体等复杂场景,显著性检测的效果并不显著,如何更好提高显著性检测效果仍是一种技术挑战。

发明内容

[0006] 本申请实施例提供了RGBD显著性检测方法以及相关装置,有利于提高显著性检测的普适性,在诸如低对比度、小物体、多个物体等复杂场景下也能取得较好的显著性检测效果。

[0007] 第一方面,本申请实施例提供了一种RGBD显著性检测方法,该方法包括:获取RGB图像和与所述RGB图像配准的深度图像;将RGB图像导入到包含N个特征提取模块和N个特征强化模块的神经网络模型进行处理,以及将深度图像导入神经网络模型进行处理,获得N个特征强化模块的N个跨模态特征图;对N个跨模态特征图进行融合,获得检测结果图。其中,N为大于等于2的整数。神经网络模型中的特征提取模块和特征强化模块交替连接设置。其中,N个特征提取模块中的第j个特征提取模块的处理包括:对所述第j个特征提取模块的输入数据进行特征提取,获得所述第j个特征提取模块的处理结果并输出到所述N个特征强化

模块中的第j个特征强化模块;其中,当j等于1时,所述输入数据为所述RGB图像,当j不等于1时,所述输入数据为第j-1个特征强化模块输出的跨模态特征图;j为大于等于1且小于等于N的整数;其中,N个特征强化模块中的第j个特征强化模块的处理包括:对导入至所述第j个特征强化模块的深度图像进行特征提取,获得深度对比度增强图;利用所述深度对比度增强图优化所述第j个特征提取模块的处理结果,获得所述第j个特征强化模块的跨模态特征图;

[0008] 可以看到,本申请实施例通过对特征提取模块和特征强化模块的交替设置,一方面,使RGB图像的信息经由N个特征提取模块中的各个特征提取模块不断进行特征提取,每经过一次特征提取都会丢失一些非显著的信息,以输出的特征图的参数和空间尺寸(尺度)。所以,通过不同的特征提取模块,可实现在不同尺度上提取RGB图像中的特征,有利于保证小物体以及复杂场景的特征显现。另一方面,特征强化模块通过对深度图像进行特征提取获得深度对比度增强图,来提高深度图像的对比度(即引入了图像边缘信息),并据此优化RGB分支的特征,从而输出对应的跨模态特征图,使得算法在低对比度场景有较好的性能。再一方面,由于输入到不同特征强化模块的RGB分支特征具有不同的尺度,所以不同特征强化模块输出的不同尺度的跨模态特征图。通过对不同尺度的跨模态特征图进行融合互补,进一步帮助优化低对比度或构图复杂应用场景的检测效果,使得RGB-D显著性检测效果更加具有普适性。

[0009] 基于第一方面,在可能的实施例中,对所述第j个特征提取模块的输入数据进行特征提取,包括:对所述第j个特征提取模块的输入数据通过至少一个卷积层处理和至少一个池化层处理。具体的,在可能的实现中,N等于5,神经网络模型包括5个交替设置的特征提取模块和特征强化模块,其中:第1个特征提取模块包括2个卷积层和1个池化层;第2个特征提取模块包括2个卷积层和1个池化层;第3个特征提取模块包括3个卷积层和1个池化层;第4个特征提取模块包括3个卷积层和1个池化层;第5个特征提取模块包括3个卷积层和1个池化层。具体的,该神经网络模型中可包括16个有模型参数的层,具体可包括13个卷积层(有模型参数)、3个全连接层(有模型参数)和5个池化层(无模型参数),示例性地,每个卷积层可均使用3*3的小型卷积核,每个池化层可均为2*2的最大池化层。这样,对于不同特征提取模块的“卷积层+池化层”组合,RGB图像的信息经由N个特征提取模块中的各个特征提取模块的卷积层和池化层不断进行特征提取。卷积层对图像进行过滤,通过卷积处理,提取某个或某些属性的特征,卷积后输出特征映射图。池化层可对该特征映射图进行池化,进一步减少参数的数量(特征图尺寸减半),得到特征向量。由于卷积和池化过程会不断地丢失一些非显著的信息,以减少特征映射图的参数和空间尺寸,越深的卷积神经网络会提取越具体明显的特征,越浅的网络提取越浅显的特征。所以,通过不同特征提取模块,可在不同尺度上提取RGB图像中的特征,即通过有效地利用VGG-16能用较低成本实现了不同尺度提取特征,有利于保证小物体以及复杂场景的特征显现。多尺度信息能帮助优化有多个物体和小物体的应用场景的检测效果。

[0010] 基于第一方面,在可能的实施例中,利用所述深度对比度增强图优化所述第j个特征提取模块的处理结果,获得所述第j个特征强化模块的跨模态特征图,包括:将深度对比度增强图与所述第j个特征提取模块的处理结果相乘,获得相乘后的结果;将相乘后的结果与所述第j个特征提取模块的处理结果相加,获得第j个特征强化模块的跨模态特征图(或

称跨模态特征图,或称跨模态特征信息)。可以看到,通过利用深度对比度增强图优化第j个特征提取模块的处理结果,使得述第j个特征强化模块输出的跨模态特征图中考虑不同尺度之间的RGB特征和深度特征的互补,有效地利用了多尺度信息以及对对比度先验信息,使得算法在低对比度场景有较好的性能,对深度图的低分辨率、噪声鲁棒性高。

[0011] 基于第一方面,在可能的实施例中,神经网络模型还包括全连接层(例如,3层的全连接层);对所述N个跨模态特征图进行融合,获得检测结果图,包括:通过金字塔模型对所述N个跨模态特征图进行融合,获得融合结果,将所述融合结果输出到所述全连接层进行处理,获得所述显著性检测的检测结果图。其中,该金字塔模型包括具有多个层的金字塔,各层对应有自己的图像集,顶层的图像集包括一个图像。底层的图像集包括N个图像,N个图像与N个特征提取模块输出的N个跨模态特征图一一对应。从金字塔的底层到金字塔的顶层的各个层的图像数量依次递减。

[0012] 在一些具体实施例中,该金字塔模型可为本申请实施例提供的第一种金字塔模型,所述第一种金字塔模型又可称为流动金字塔模型,在利用流动金字塔模型进行数据融合过程中,按照从所述金字塔的底层到所述金字塔的顶层的顺序,获得除底层外的每一层的图像集;其中,除底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度不大于所述任一图像的尺度的所有图像融合而成,不同图像融合的方式具体为像素级别的相加。这样,各个尺度的跨模态特征进入流动金字塔后,逐层相互作用至金字塔的顶层,实现了充分的互补和融合,金字塔的顶层的融合图像即为最终得到融合结果。可以看到,本申请实施例提供的流动金字塔引入了更多的跨尺度连接来获得特征融合结果,使得多尺度的跨模态特征融合与互补更为充分。对于任一尺度特征而言,其拥有了更多尺度的RGB特征的信息,不同尺度之间的跨模态特征融合,促进了RGB特征与深度特征更好地互补。所以本申请实施例引入流动金字塔能够保证其RGB-D图像数据的兼容性和互补性,进一步帮助优化低对比度或构图复杂应用场景的检测效果,使得RGB-D显著性检测效果更加具有普适性。

[0013] 在又一些具体实施例中,该金字塔模型也可以是本申请实施例提供的第二种金字塔模型,在利用第二种金字塔模型进行数据融合过程中,按照从所述金字塔的底层到所述金字塔的顶层的顺序,获得除底层外的每一层的图像集;其中,除底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度同所述任一图像相同的一个图像和尺度比所述任一图像小的一个图像两者融合而成,不同图像融合的方式具体为像素级别的相加。第二种金字塔模型例如可以是传统的金字塔模型,这种情况下,除底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度同所述任一图像相同的一个图像和尺度比所述任一图像小一级的图像两者融合而成。各个尺度的跨模态特征进入第二种金字塔模型后,逐层相互作用至金字塔的顶层,也能在一定程度上实现了互补和融合,金字塔的顶层的融合图像即为最终得到融合结果。

[0014] 基于第一方面,在可能的实施例中,对导入至所述第j个特征强化模块的深度图像进行特征提取,获得深度对比度增强图,包括:对导入至所述第j个特征强化模块的深度图像通过卷积层集合处理以及通过激活函数处理,获得深度对比度增强图。其中,所述卷积层集合中各个卷积层的卷积核参数是利用对比度增强损失函数对导入至所述第j个特征强化模块的深度图像进行监督而获得的。

[0015] 基于第一方面,在可能的实施例中,利用对比度增强损失函数对导入至所述第j个

特征强化模块的深度图像进行监督,包括:将导入至所述第j个特征强化模块的深度图像通过所述卷积层集合的各个卷积层进行卷积处理,获得深度特征图;利用对比度增强损失函数对所述深度特征图进行运算,得到所述对比度增强损失函数的值;在所述对比度增强损失函数的值未达到局部最小值的情况下,更新所述卷积层集合的各个卷积层的卷积核参数,继续梯度下降迭代,直到达到局部最小值。使用本申请实施例提供的对比度增强损失函数对深度特征进行监督,能够提高深度图的对比度,引入了图像边缘信息,获得深度对比度增强图。将该深度对比度增强图与从RGB特征图中的特征相乘,最后将相乘后的结果与RGB特征图相加,可得到跨模态特征。

[0016] 基于第一方面,在可能的实施例中,对比度增强损失函数 l_c 的数学描述如下所示:

$$[0017] \quad l_c = a_1 l_f + a_2 l_b + a_3 l_w$$

[0018] 其中,

$$[0019] \quad l_f = -\log\left(1 - 4 \times \sum_{p \in F} \frac{(p - \widehat{p}_f)^2}{N_f}\right)$$

$$[0020] \quad l_b = -\log\left(1 - 4 \times \sum_{p \in B} \frac{(p - \widehat{p}_b)^2}{N_b}\right)$$

$$[0021] \quad l_w = -\log(\widehat{p}_f - \widehat{p}_b)^2$$

$$[0022] \quad \widehat{p}_f = \sum_{p \in F} \frac{p}{N_f}$$

$$[0023] \quad \widehat{p}_b = \sum_{p \in B} \frac{p}{N_b}$$

[0024] 其中, a_1 、 a_2 和 a_3 分别设置为5、5和1; l_f 表示监督深度图像的前景部分的训练; F 为前景点集合, \widehat{p}_f 为前景点均值, N_f 为前景点个数; l_b 表示监督深度图像的背景部分的训练; B 为背景点集合, \widehat{p}_b 为背景点均值, N_b 为背景点个数; l_w 表示监督深度图像的全图训练; \widehat{p}_f 为前景点均值, \widehat{p}_b 为背景点均值。

[0025] 可以看到,本申请实施例通过FEM模块的对比度增强损失函数对深度特征进行监督,提取深度图中的对比度先验,从而提高深度图的对比(引入了图像边缘信息),并据此优化RGB分支的特征,使得算法在低对比度场景有较好的性能,对深度图的低分辨率、噪声鲁棒性高。通过考虑不同尺度之间的RGB特征和深度特征的互补,有效地利用了多尺度信息以及对比度先验信息,从而能帮助优化低对比度或构图复杂应用场景的检测效果。

[0026] 基于第一方面,在可能的实施例中,经金字塔融合处理的数据可被输入到全连接层,本申请实施例中的全连接层的数量可以是3个,各个全连接层的神经元数量例如可以为4096、4096和1000。通过全连接层对数据进行整合和降维,从而得到与显著性分类总数相同的归一化的检测结果图。最后将检测结果图输出,从而实现了RGB-D图像数据的显著性检测。

[0027] 基于第一方面,在可能的实施例中,对检测结果图后续还可以应用于图像分割,目标识别,视觉跟踪,图像检索,图像编辑,图像语义分割,虚拟大光圈拍照,翻译,文本分析等等应用场景中,这里不做限定。

[0028] 第二方面,本申请实施例提供了一种用于RGBD显著性检测的神经网络系统,该神

神经网络系统包括神经网络模型和融合模块,所述神经网络模型包括N个特征提取模块和N个特征强化模块,所述特征提取模块和所述特征强化模块交替连接;每个特征提取模块包括至少一个卷积层和至少一个池化层;可选的,所述神经网络模型还包括全连接层,其中:神经网络模型用于,对获取的RGB图像进行处理,以及对与所述RGB图像配准的深度图像进行处理,获得所述N个特征强化模块的N个跨模态特征图;N为大于等于2的整数;其中,所述N个特征提取模块中的第j个特征提取模块具体用于:对所述第j个特征提取模块的输入数据进行特征提取,获得所述第j个特征提取模块的处理结果并输出到所述N个特征强化模块中的第j个特征强化模块;其中,当j等于1时,所述输入数据为所述RGB图像,当j不等于1时,所述输入数据为第j-1个特征强化模块输出的跨模态特征图;j为大于等于1且小于等于N的整数;其中,所述N个特征强化模块中的第j个特征强化模块具体用于:对导入至所述第j个特征强化模块的深度图像进行特征提取,获得深度对比度增强图;利用所述深度对比度增强图优化所述第j个特征提取模块的处理结果,获得所述第j个特征强化模块的跨模态特征图;融合模块用于,对所述N个跨模态特征图进行融合,获得显著性检测的检测结果图。所述融合模块例如包括本申请实施例提供的第一金字塔模型(流动金字塔模型),所述融合模块又例如包括本申请实施例提供的第二金字塔模型(例如,传统金字塔模型)。

[0029] 可以看到,本申请实施例提供的神经网络系统,一方面,能使RGB图像的信息经由N个特征提取模块中的各个特征提取模块不断进行特征提取,每经过一次特征提取都会丢失一些非显著的信息,以输出的特征图的参数和空间尺寸(尺度)。所以,通过不同的特征提取模块,可实现在不同尺度上提取RGB图像中的特征,有利于保证小物体以及复杂场景的特征显现。另一方面,利用特征强化模块通过对深度图像进行特征提取获得深度对比度增强图,来提高深度图像的对比度(即引入了图像边缘信息),并据此优化RGB分支的特征,从而输出对应的跨模态特征图,使得算法在低对比度场景有较好的性能。再一方面,不同特征强化模块输出的不同尺度的跨模态特征图,通过融合模块对不同尺度的跨模态特征图进行融合互补,进一步帮助优化低对比度或构图复杂应用场景的检测效果,使得RGB-D显著性检测效果更加具有普适性。

[0030] 基于第二方面,在可能的实施例中,N等于5,神经网络系统包括5个特征提取模块和5个特征强化模块,其中:第1个特征提取模块包括2个卷积层和1个池化层;第2个特征提取模块包括2个卷积层和1个池化层;第3个特征提取模块包括3个卷积层和1个池化层;第4个特征提取模块包括3个卷积层和1个池化层;第5个特征提取模块包括3个卷积层和1个池化层。

[0031] 基于第二方面,在可能的实施例中,第j个特征提取模块具体用于:对所述第j个特征提取模块的输入数据通过至少一个卷积层处理和至少一个池化层处理。

[0032] 基于第二方面,在可能的实施例中,第j个特征强化模块具体用于:将所述深度对比度增强图与所述第j个特征提取模块的处理结果相乘,获得相乘后的结果;将所述相乘后的结果与所述第j个特征提取模块的处理结果相加,获得所述第j个特征强化模块的跨模态特征图。

[0033] 基于第二方面,在可能的实施例中,神经网络模型还包括全连接层;所述融合模块具体用于:通过金字塔模型对所述N个跨模态特征图进行融合,获得融合结果;将所述融合结果输出到所述全连接层进行处理,获得所述显著性检测的检测结果图。

[0034] 基于第二方面,在可能的实施例中,金字塔模型包括具有多个层的金字塔,从所述金字塔的底层到所述金字塔的顶层的各个层的图像数量依次递减,所述底层的图像集包括N个图像,所述N个图像与所述N个特征提取模块输出的N个跨模态特征图一一对应;所述融合模块具体用于:按照从所述金字塔的底层到所述金字塔的顶层的顺序,获得除所述底层外的每一层的图像集;其中,除所述底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度不大于所述任一图像的尺度的所有图像融合而成,所述顶层的图像集包括一个图像;将所述顶层的所述一个图像作为所述融合结果。

[0035] 基于第二方面,在可能的实施例中,金字塔模型包括具有多个层的金字塔,从所述金字塔的底层到所述金字塔的顶层的各个层的图像数量依次递减,所述底层的图像集包括N个图像,所述N个图像与所述N个特征提取模块输出的N个跨模态特征图一一对应;所述融合模块具体用于:按照从所述金字塔的底层到所述金字塔的顶层的顺序,通过融合填充获得除所述底层外的每一层的图像集;其中,除所述底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度同所述任一图像相同的一个图像和尺度比所述任一图像小的一个图像两者融合而成,所述顶层的图像集包括一个图像;将所述顶层的所述一个图像作为所述融合结果。

[0036] 基于第二方面,在可能的实施例中,第j个特征强化模块具体用于:对导入至所述第j个特征强化模块的深度图像通过卷积层集合处理以及通过激活函数处理,获得深度对比度增强图。其中,卷积层集合中各个卷积层的卷积核参数是所述第j个特征强化模块利用对比度增强损失函数对导入至所述第j个特征强化模块的深度图像进行监督而获得的。

[0037] 基于第二方面,在可能的实施例中,第j个特征强化模块具体用于:将导入至所述第j个特征强化模块的深度图像通过所述卷积层集合的各个卷积层进行卷积处理,获得深度特征图;利用对比度增强损失函数对所述深度特征图进行运算,得到所述对比度增强损失函数的值;在所述对比度增强损失函数的值未达到局部最小值的情况下,更新所述卷积层集合的各个卷积层的卷积核参数,继续梯度下降迭代,直到达到局部最小值。

[0038] 神经网络系统的各模块的具体可用于实现第一方面的各实施例所描述的方法。具体内容可参考第一方面的各实施例的相关描述。

[0039] 第三方面,本申请实施例提供一种计算节点(计算装置),其特征在于,所述计算节点(计算装置)包括处理器以及耦合的存储器;其中:所述存储器用于存储RGB图像和深度图像,还用于存储程序指令;所述处理器用于调用所述程序指令,以实现如第一方面的任意实施例所描述的方法。

[0040] 第四方面,本申请实施例提供了一种计算节点集群,包括:至少一个计算节点,每个计算节点包括处理器和存储器,所述处理器执行所述存储器中的代码执行如第一方面任一项所述的方法。

[0041] 第五方面,本申请实施例提供一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行如第一方面的任意实施例所描述的方法。

[0042] 第六方面,提供了一种计算机程序产品,当所述计算机程序产品被计算机读取并执行时,如第一方面任一项所述的方法将被执行。

[0043] 可以看到,本申请实施例通过实现了不同尺度提取特征,有利于保证小物体以及

复杂场景的特征显现。通过特征强化模块增加深度图的对比度先验,提高了深度图的对比度,得到更为清晰的目标边缘,能够优化低对比度或构图复杂应用场景的检测效果。通过多尺度的跨模态特征融合,更好地实现了RGB信息和深度信息的跨模态互补,保证其RGB-D图像数据的兼容性和互补性。所以,实施本申请实施例能够实现更为准确地检测显著性物体,有利于提高显著性检测的普适性,在诸如低对比度、小物体、多个物体等复杂场景下也能取得较好的显著性检测效果。

附图说明

[0044] 为了更清楚地说明本申请实施例或背景技术中的技术方案,下面将对本申请实施例或背景技术中所需要使用的附图进行说明。

[0045] 图1为本申请实施例提供的一种人工智能主体框架示意图;

[0046] 图2a为本申请实施例提供的一种系统架构示意图;

[0047] 图2b为本申请实施例提供的又一种系统架构示意图;

[0048] 图3为本申请实施例提供的一种卷积神经网络结构示意图;

[0049] 图4为现有的一种神经网络结构的示意图;

[0050] 图5为本申请实施例提供的一种神经网络结构(神经网络系统)的示意图;

[0051] 图6为本申请实施例提供的一种特征强化模块的数据处理流程示意图;

[0052] 图7为本申请实施例提供的又一种神经网络结构(神经网络系统)的示意图;

[0053] 图8为本申请实施例提供的针对各种设计场景的实验效果图;

[0054] 图9为本申请实施例提供的一种芯片硬件结构图;

[0055] 图10为本申请实施例提供的一种装置的结构示意图;

[0056] 图11为本申请实施例提供的一种多处理器系统的结构示意图;

[0057] 图12为本申请实施例提供的一种计算节点(计算装置)的结构示意图;

[0058] 图13为本申请实施例提供的RGBD显著性检测方法的流程示意图;

[0059] 图14为本申请实施例提供的针对各种图像场景的实验效果图。

具体实施方式

[0060] 下面结合本申请实施例中的附图对本申请实施例进行描述。本申请的实施方式部分使用的术语仅用于对本申请的具体实施例进行解释,而非旨在限定本申请。

[0061] 首先描述本文中涉及的一些概念。

[0062] (1) RGB-D图像数据。在3D计算机图形中,RGB-D图像数据包括两幅图像:一个是普通的RGB三通道彩色图像(简称RGB图像),另一个是深度图像(Depth Map)。

[0063] RGB图像中的像素值是一个红绿蓝(BlueGreenRed,RGB)颜色值,像素值可以是表示颜色的长整数。例如,像素值为 $256 * R + 100 * G + 76B$,其中,R代表红色分量,G代表绿色分量,B代表蓝色分量。各个颜色分量中,数值越小,亮度越低,数值越大,亮度越高。对于灰度图像来说,像素值可以是灰度值。

[0064] 深度图像是包含与视点的场景对象的表面的距离有关的信息的图像或图像通道。深度图像类似于灰度图像,只是它的每个像素值是传感器距离物体的实际距离。

[0065] RGB图像和深度图像通常是配准的,因而像素点之间具有一对一的对应关系。

[0066] (2) 图像特征。图像的特征提取的主要目的是对图像的大量像素点数据进行降维。特征提取的主要思想是将原始图像样本投影到一个低维特征空间,得到最能反应图像样本本质或进行样本区分的低维样本特征,这样的特征又可称为图像特征。通常来说,图像特征主要有图像的颜色特征、纹理特征、形状特征和空间关系特征等。其中:

[0067] 颜色特征是一种全局特征,描述了图像或图像区域所对应的景物的表面性质;一般颜色特征是基于像素点的特征,此时所有属于图像或图像区域的像素都有各自的贡献。由于颜色对图像或图像区域的方向、大小等变化不敏感,所以颜色特征不能很好地捕捉图像中对象的局部特征。

[0068] 纹理特征也是一种全局特征,它也描述了图像或图像区域所对应景物的表面性质;但由于纹理只是一种物体表面的特性,并不能完全反映出物体的本质属性,所以仅仅利用纹理特征是无法获得高层次图像内容的。与颜色特征不同,纹理特征不是基于像素点的特征,它需要在包含多个像素点的区域中进行统计计算。

[0069] 形状特征有两类表示方法,一类是轮廓特征,另一类是区域特征,图像的轮廓特征主要针对物体的外边界,而图像的区域特征则关系到整个形状区域;

[0070] 空间关系特征,是指图像中分割出来的多个目标之间的相互的空间位置或相对方向关系,这些关系也可分为连接/邻接关系、交叠/重叠关系和包含/包容关系等。通常空间位置信息可以分为两类:相对空间位置信息和绝对空间位置信息。前一种关系强调的是目标之间的相对情况,如上下左右关系等,后一种关系强调的是目标之间的距离大小以及方位。

[0071] 需要说明的,上述列举的图像特征可以作为图像中具有的特征的一些举例,图像还可以具有其他特征,如更高层级的特征:语义特征,此处不再展开。

[0072] 请参见图1,图1示出本申请实施例相关的一种人工智能主体框架示意图,该主体框架描述了人工智能系统总体工作流程,适用于通用的人工智能领域需求。

[0073] 下面从“智能信息链”(水平轴)和“IT价值链”(垂直轴)两个维度对上述人工智能主题框架进行阐述。

[0074] “智能信息链”反映从数据的获取到处理的一系列过程。举例来说,可以是智能信息感知、智能信息表示与形成、智能推理、智能决策、智能执行与输出的一般过程。在这个过程中,数据经历了“数据—信息—知识—智慧”的凝练过程。

[0075] “IT价值链”从人工智能的底层基础设施、信息(提供和处理技术实现)到系统的产业生态过程,反映人工智能为信息技术产业带来的价值。下面进行进一步描述相关概念。

[0076] (1) 基础设施:基础设施为人工智能系统提供计算能力支持,实现与外部世界的沟通,并通过基础平台实现支撑。基础设施例如包括传感器、智能芯片、基础平台等。具体的,传感器可用于与外部沟通;智能芯片(CPU、NPU、GPU、ASIC、FPGA等硬件加速芯片)可用于提供计算能力;基础平台包括分布式计算框架及网络等相关的平台保障和支持,可进一步包括云存储和计算、互联互通网络等。举例来说,传感器和外部沟通获取数据,这些数据提供给基础平台提供的分布式计算系统中的智能芯片进行计算。

[0077] (2) 数据:基础设施的上一层的数据用于表示人工智能领域的数据来源。数据例如可以是图形、图像、视频、音频、文本等等。

[0078] (3) 数据处理:数据处理通常包括数据训练、机器学习、深度学习,以及搜索、推理、

决策等方式。其中,机器学习和深度学习可以对数据进行符号化和形式化的智能信息建模、抽取、预处理、训练等。推理是指在计算机或智能系统中,模拟人类的智能推理方式,依据推理控制策略,利用形式化的信息进行机器思维和求解问题的过程,典型的功能是搜索与匹配。决策是指智能信息经过推理后进行决策的过程,通常提供分类、排序、预测等功能。

[0079] (4) 通用能力:对数据经过上面提到的数据处理后,进一步基于数据处理的结果可以形成一些通用的能力,比如可以是算法或者一个通用系统,例如,翻译,文本的分析,计算机视觉的处理,语音的识别,视频图像的识别与处理等等。

[0080] 具体的,本申请实施例中提出的显著性物体检测方法可具体应用于下述应用场景:图像分割(image segmentation),目标识别(object recognition),视觉跟踪(visual tracking),图像检索(image retrieval),图像编辑(image editing),图像语义分割(semantic segmentation),等等。其中:

[0081] 图像分割是指将数字图像细分为多个图像子区域(像素的集合,也被称作超像素)的过程。图像分割的目的是简化或改变图像的代表形式,使得图像更容易理解和分析。

[0082] 目标识别是指一个特殊目标(或一种类型的目标)从其它目标(或其它类型的目标)中被区分出来的过程。它既包括两个非常相似目标的识别,也包括一种类型的目标同其他类型目标的识别。

[0083] 视觉跟踪是指对图像序列中的运动目标进行检测、提取、识别和跟踪,获得运动目标的运动参数,如位置、速度、加速度和运动轨迹等,从而进行下一步的处理与分析,实现对运动目标的行为理解,以完成更高一级的检测任务。

[0084] 图像检索是指用户向图像检索系统输入文本或图片,从图像数据库中检索出相同类别的图像。从图像数据库的图像中检测出显著区域,也就是目标物所在区域,对该区域进行特征提取、分析以及标注,可以提高图像检索的准确率。另一方面,对基于内容的图像检索任务,检测出显著区域,更有利于针对性地提取特征以及检索。

[0085] 图像编辑是指改变图像的过程,图像包括数码照片,传统的模拟照片和插图。对模拟图像的编辑通常也称为照片修饰,使用工具为喷枪等。数码照片的编辑可采用图形处理软件,在检测出显著区域后,有利于对特定区域的自动、半自动编辑(比如改变头发颜色)。

[0086] 图像语义分割是在像素级别上的分类,属于同一类的像素都要被归为一类。在图像领域,语义指的是图像的内容,对图片意思的理解;分割的意思是从像素的角度分割出图片中的不同对象,对原图中的各类像素进行标注或说明。

[0087] 本申请实施例所描述的显著目标检测可在上述应用场景中可取得良好的效果。值得注意的是,本申请应用场景不限于以上场景,其他与显著目标检测相关的场景,通过采集对应的RGB-D图像,经过本申请的神经网络训练流程,也可以达到较好的检测结果。

[0088] (5) 智能产品及行业应用:智能产品及行业应用指人工智能系统在各领域的产品和应用,是对人工智能整体解决方案的封装,将智能信息决策产品化、实现落地应用,其应用领域主要包括:智能制造、智能交通、智能家居、智能医疗、智能安防、自动驾驶,平安城市,智能终端等。

[0089] 参见图2a,本申请实施例提供了一种系统架构200。如图2a所示,系统架构200可包括执行设备210,训练设备220,数据库230,客户设备240,数据存储系统250和数据采集设备260,其中:执行设备210可分别与训练设备220、客户设备240、数据采集设备260通信连接,

数据库230分别与数据采集设备260、训练设备220通信连接。可能的实施例中,数据库230还可与客户设备240通信连接。

[0090] 数据采集设备260用于采集样本数据(例如,大量用于训练的RGB-D图像数据)并存入数据库230,训练设备220基于数据库230中维护的样本数据生成目标模型201。下面将更详细地描述训练设备220如何基于样本数据得到目标模型201,目标模型201例如为本申请实施例描述的神经网络模型。

[0091] 本申请实施例描述的神经网络模型可以是由神经单元组成的,神经单元可以是指以 x_s 和截距1为输入的运算单元,该运算单元的输出可以为:

$$[0092] \quad h_{w,b}(x) = f(W^T x) = f\left(\sum_{s=1}^n W_s x_s + b\right) \quad (1-1)$$

[0093] 其中, $s=1,2,\dots,n$, n 为大于1的自然数, W_s 为 x_s 的权重, b 为神经单元的偏置。 f 为神经单元的激活函数(activation functions),用于将非线性特性引入神经网络中,来将神经单元中的输入信号转换为输出信号。该激活函数的输出信号可以作为下一层的输入。激活函数可以是sigmoid函数。神经网络是将许多个上述单一的神经单元联结在一起形成的网络,即一个神经单元的输出可以是另一个神经单元的输入。每个神经单元的输入可以与前一层的局部接受域相连,来提取局部接受域的特征,局部接受域可以是由若干个神经单元组成的区域。

[0094] 训练神经网络的目的,就是最终得到训练好的神经网络的各个层的相关参数(如由很多层的向量 W 形成的权重矩阵)。因此,神经网络的训练过程本质上就是学习控制空间变换的方式,更具体的就是学习权重矩阵。

[0095] 因为希望神经网络的输出尽可能的接近真正想要预测的值,所以可以通过比较当前网络的预测值和真正想要的目标值,再根据两者之间的差异情况来更新每一层神经网络的权重向量(当然,在第一次更新之前通常会有初始化的过程,即为神经网络中的各层预先配置参数),比如,如果网络的预测值高了,就调整权重向量让它预测低一些,不断的调整,直到神经网络能够预测出真正想要的目标值。因此,就需要预先定义“如何比较预测值和目标值之间的差异”,这便是损失函数(loss function)或目标函数(objective function),它们是用于衡量预测值和目标值的差异的重要方程。其中,以损失函数举例,损失函数的输出值(lossvalue)越高表示差异越大,那么神经网络的训练就变成了尽可能缩小这个loss的过程。

[0096] 训练设备220训练得到的目标模型201可以应用不同的系统或设备中。在附图2a中可应用到执行设备210。

[0097] 具体的,执行设备210可配置有I/O接口212,与外部设备进行数据交互,“用户”可以通过客户设备240向I/O接口212输入数据(如RGB-D图像数据)。

[0098] 执行设备210可以调用数据存储系统250中的数据(如RGB-D图像数据)、代码等,也可以将数据(如RGB-D图像数据)、指令等存入数据存储系统250中。

[0099] 计算模块211使用目标模型201对输入的数据进行处理。本申请实施例中,计算模块211可利用目标模型201实现将RGB图像和深度图像融合为多尺度跨模态特征,进而通过金字塔结构的互补/融合处理获得显著性检测的结果图(可简称检测结果图)。

[0100] 在可能的实现中,还可在执行设备210中设置关联功能模块213,关联功能模块213

对检测结果图进行后续加工处理,例如可将检测结果图应用于图像翻译,图像文本分析,图像分割,目标识别,视觉跟踪,图像检索,图像编辑,图像语义分割等等应用场景中。

[0101] 最后,I/O接口212将处理结果返回给客户设备240,提供给用户。

[0102] 更深层地,训练设备220可以针对不同的目标,基于不同的数据生成相应的目标模型201,以给用户提供更佳的结果。

[0103] 在图2a中所示情况下,用户可以手动指定输入执行设备210中的数据,例如,在I/O接口212提供的界面中操作。另一种情况下,客户设备240可以自动地向I/O接口212输入数据并获得结果,如果客户设备240自动输入数据需要获得用户的授权,用户可以在客户设备240中设置相应权限。用户可以在客户设备240查看执行设备210输出的结果,具体的呈现形式可以是显示、声音、动作等具体方式。在可能的实现中,客户设备240也可以作为数据采集端将采集到的样本数据(例如,经过人工标注的RGB-D图像数据)存入数据库230。

[0104] 值得注意的,图2a中所示设备、器件、模块等之间的位置关系不构成任何限制,例如,在图2a中,数据存储系统250相对执行设备210是外部存储器,在其它情况下,也可以将数据存储系统250置于执行设备210中。

[0105] 参见图2b,图2b是本申请实施例提供的又一种系统架构300。系统架构300可包括一个或多个执行设备210,数据存储系统250,以及本地设备(如图示中的本地设备301和本地设备302),其中:

[0106] 执行设备210由一个或多个服务器实现,可选的,与其它计算设备配合,例如:数据存储、路由器、负载均衡器等设备;执行设备210可以布置在一个物理站点上,或者分布在多个物理站点上。执行设备210可以使用数据存储系统250中的数据(例如,RGB-D图像数据),或者调用数据存储系统250中的程序代码实现本申请实施例所描述的显著性检测方法。

[0107] 用户可以操作各自的用户设备(例如本地设备301和本地设备302)与执行设备210进行交互。每个本地设备可以表示任何计算设备,例如个人计算机、计算机工作站、智能手机、平板电脑、智能摄像头、智能汽车或其他类型蜂窝电话、媒体消费设备、可穿戴设备、机顶盒、游戏机等。

[0108] 每个用户的本地设备可以通过任何通信机制/通信标准的通信网络与执行设备210进行交互,通信网络可以是广域网、局域网、点对点连接等方式,或它们的任意组合。

[0109] 在另一种实现中,执行设备210的一个方面或多个方面可以由每个本地设备实现,例如,本地设备301可以为执行设备210提供本地数据或反馈计算结果。

[0110] 需要注意的,执行设备210的所有功能也可以由本地设备实现。例如,本地设备301实现执行设备210的功能并为自己的用户提供服务,或者为本地设备302的用户提供服务。

[0111] 本申请实施例中描述的神经网络模型以卷积神经网络(Convolutional Neural Network, CNN)为基础构型,下面对卷积神经网络进行描述。

[0112] 卷积神经网络是一种带有卷积结构的深度神经网络,可以是一种深度学习(deep learning)架构,深度学习架构是指通过机器学习的算法,在不同的抽象层级上进行多个层次的学习。作为一种深度学习架构,CNN是一种前馈(feed-forward)人工神经网络,该前馈人工神经网络中的各个神经元对输入其中的图像中的重叠区域作出响应。

[0113] 卷积神经网络包含了一个由卷积层和子采样层构成的特征抽取器。该特征抽取器

可以看作是滤波器,卷积过程可以看作是使用一个可训练的滤波器与一个输入的图像或者卷积特征平面(feature map)做卷积。卷积层是指卷积神经网络中对输入信号进行卷积处理的神经元层。在卷积神经网络的卷积层中,一个神经元可以只与部分邻层神经元连接。一个卷积层中,通常包含若干个特征平面,每个特征平面可以由一些矩形排列的神经单元组成。同一特征平面的神经单元共享权重,这里共享的权重就是卷积核。共享权重可以理解成提取图像信息的方式与位置无关。这其中隐含的原理是:图像的某一部分的统计信息与其他部分是一样的。即意味着在某一部分学习的图像信息也能用在另一部分上。所以对于图像上的所有位置,我们都能使用同样的学习得到的图像信息。在同一卷积层中,可以使用多个卷积核来提取不同的图像信息,一般地,卷积核数量越多,卷积操作反映的图像信息越丰富。卷积核可以以随机大小的矩阵的形式初始化,在卷积神经网络的训练过程中卷积核可以通过学习得到合理的权重。另外,共享权重带来的直接好处是减少卷积神经网络各层之间的连接,同时又降低了过拟合的风险。

[0114] 卷积神经网络可以采用误差反向传播(back propagation, BP)算法在训练过程中修正初始的超分辨率模型中参数的大小,使得超分辨率模型的重建误差损失越来越小。具体地,前向传递输入信号直至输出会产生误差损失,通过反向传播误差损失信息来更新初始的超分辨率模型中参数,从而使误差损失收敛。反向传播算法是以误差损失为主导的反向传播运动,旨在得到最优的超分辨率模型的参数,例如权重矩阵。

[0115] 请参见图3,图3为本申请实施例涉及的一种卷积神经网络(CNN)100的结构示意图。如图3所示,卷积神经网络(CNN)100可以包括输入层110,卷积层/池化层120,以及神经网络层130。

[0116] 输入层110可以处理多维数据,例如可获取并处理RGB图像,获取并处理与RGB图像配准的深度图像;常见地,一维卷积神经网络的输入层接收一维或二维数组,其中一维数组通常为时间或频谱采样;二维数组可能包含多个通道;二维卷积神经网络的输入层接收二维或三维数组;三维卷积神经网络的输入层接收四维数组。具体的,本申请实施例中输入层110可设置RGB数据通道。

[0117] 由于使用梯度下降进行学习,卷积神经网络的输入特征可进行标准化处理。具体地,在将学习数据输入卷积神经网络前,需在通道或时间/频率维对输入数据进行归一化处理。输入特征的标准化有利于提升算法的运行效率和学习表现。

[0118] 卷积层/池化层120可以包括如示例121-126层,在一种实现中,121层为卷积层,122层为池化层,123层为卷积层,124层为池化层,125为卷积层,126为池化层;在另一种实现方式中,121、122为卷积层,123为池化层,124、125为卷积层,126为池化层。即卷积层的输出可以作为随后的池化层的输入,也可以作为另一个卷积层的输入以继续进行卷积操作。

[0119] 以卷积层121为例,卷积层121可以包括很多个卷积算子,卷积算子也称为卷积核,其在图像处理中的作用相当于一个从输入图像矩阵中提取特定信息的过滤器,卷积算子本质上可以是一个权重矩阵,这个权重矩阵通常被预先定义,在对图像进行卷积操作的过程中,权重矩阵通常在输入图像上沿着水平方向一个像素接着一个像素(或两个像素接着两个像素……这取决于步长stride的取值)的进行处理,从而完成从图像中提取特定特征的工作。该权重矩阵的大小应该与图像的大小相关,需要注意的是,权重矩阵的纵深维度(depth dimension)和输入图像的纵深维度是相同的,在进行卷积运算的过程中,权重矩阵

会延伸到输入图像的整个深度。因此,和一个单一的权重矩阵进行卷积会产生一个单一纵深维度的卷积化输出,但是大多数情况下不使用单一权重矩阵,而是应用维度相同的多个权重矩阵。每个权重矩阵的输出被堆叠起来形成卷积图像的纵深维度。不同的权重矩阵可以用来提取图像中不同的特征,例如一个权重矩阵用来提取图像边缘信息,另一个权重矩阵用来提取图像的特定颜色,又一个权重矩阵用来对图像中不需要的噪点进行模糊化……该多个权重矩阵维度相同,经过该多个维度相同的权重矩阵提取后的特征图维度也相同,再将提取到的多个维度相同的特征图合并形成卷积运算的输出。

[0120] 这些权重矩阵中的权重值在实际应用中需要经过大量的训练得到,通过训练得到的权重值形成的各个权重矩阵可以从输入图像中提取信息,从而帮助卷积神经网络100进行正确的预测。

[0121] 需要说明的是,上述121—126层仅仅为示例,实际上还可设置更多的卷积层和/或更多的池化层。当卷积神经网络100有多个卷积层的时候,初始的卷积层(例如121)往往提取较多的一般特征,该一般特征也可以称之为低级别的特征;随着卷积神经网络100深度的加深,越往后的卷积层(例如126)提取到的特征越来越复杂,比如高级别的语义之类的特征。本申请实施例利用了不同尺度的特征以辅助解决相关技术问题。

[0122] 由于常常需要减少训练参数的数量,因此卷积层之后常常需要周期性的引入池化层,即如图3中120所示例的121—126各层,可以是一层卷积层后面跟一层池化层,也可以是多层卷积层后面接一层或多层池化层。在图像处理过程中,池化层可用来减少图像的空间大小。池化层可以包括平均池化算子和/或最大池化算子,以用于对输入图像进行采样得到较小尺寸的图像。平均池化算子可以在特定范围内对图像中的像素值进行计算产生平均值。最大池化算子可以在特定范围内取该范围内值最大的像素作为最大池化的结果。另外,就像卷积层中用权重矩阵的大小应该与图像大小相关一样,池化层中的运算符也应该与图像的大小相关。通过池化层处理后输出的图像尺寸可以小于输入池化层的图像的尺寸,池化层输出的图像中每个像素点表示输入池化层的图像的对应子区域的平均值或最大值。

[0123] 通常来讲,卷积层中的卷积核包含权重系数(权重矩阵),而池化层不包含权重系数,因此在一些场景中,也可将池化层不视为独立的层。

[0124] 在经过卷积层/池化层120的处理后,卷积神经网络100还不足以输出所需要的输出信息。因为如前所述,卷积层/池化层120只会提取特征,并减少输入图像带来的参数。然而为了生成最终的输出信息(所需要的类信息或别的相关信息),卷积神经网络100需要利用神经网络层130来生成一个或者一组所需要的类的数量的输出。因此,在神经网络层130中可以包括多层隐含层(如图3所示的131、132至13n)以及输出层140,该多层隐含层中所包含的参数可以根据具体的任务类型的相关训练数据进行预先训练得到,例如该任务类型可以包括图像识别,图像分类,图像超分辨率重建等等……

[0125] 卷积神经网络中的隐含层例如包括全连接(fully-connected, FC)层,全连接层通常只向其它全连接层传递信号。特征图在全连接层中会失去3维结构,被展开为向量并通过激励函数传递至下一层。在一些可能卷积神经网络中,全连接层的功能可部分由全局均值池化(global average pooling)取代,全局均值池化会将特征图每个通道的所有值取平均。

[0126] 在神经网络层130中的多层隐含层之后,也就是整个卷积神经网络100的最后层为

输出层140,该输出层140具有类似分类交叉熵的损失函数,具体用于计算预测误差,一旦整个卷积神经网络100的前向传播(如图3由110至140的传播为前向传播)完成,反向传播(如图3由140至110的传播为反向传播)就会开始更新前面提到的各层的权重值以及偏差,以减少卷积神经网络100的损失及卷积神经网络100通过输出层输出的结果和理想结果之间的误差。

[0127] 输出层140可使用逻辑函数或归一化指数函数(softmax function)输出分类标签。例如在显著性检测中,输出层可设计为输出图像的检测结果图(如图像中的显著目标);又例如显著性检测的结果应用到物体识别(object detection)中时,输出层可设计为输出物体的中心坐标、大小和分类等。又例如显著性检测的结果应用到图像语义分割中时,输出层可直接输出各个像素的分类结果,等等。

[0128] 需要说明的是,如图3所示的卷积神经网络100仅作为一种卷积神经网络的示例,在具体的应用中,卷积神经网络还可以以其他网络模型的形式存在,例如,将多个卷积层/池化层并行,将分别提取的特征均输入给神经网络层130进行处理。

[0129] 通常来说,RGB图像(二维图像)能够体现被拍摄物理的颜色、光照、纹理等特性,但又能提供详细的灰度信息。深度图像(三维图像)能够体现被拍摄物体的三维信息,并且不受颜色变化以及光照等的影响。因此,将二维图像和三维图像有效地结合起来,可以提供拍摄场景更加丰富全面的信息,通常可称之为多模态信息。随着深度学习的发展,卷积神经网络被应用于RGBD显著性检测领域,通过将卷积神经网络模型在大体量数据集上训练以提高显著性检测的效果。但是,由于RGB-D图像数据同时包括了RGB图像和深度图像,所以在卷积神经网络训练过程中就涉及到了如何对这两种图像的特征进行融合的问题。

[0130] 参见图4,图4示出了一种现有的神经网络结构进行RGB-D图像融合的方案。该方法从两个支路分别输入RGB图像和深度图像,输出对应的图像边缘检测图。具体的,该神经网络结构包含两个数据处理分支,分别为对应RGB图像分支和深度图像分支,每个分支都在五个尺度下进行卷积操作,提取图像特征。每个尺度下通过CA-Fuse(complementarity-aware fusion,互补融合)模块独立完成跨模态特征的融合,具体的,CA-Fuse通过行跨模态的残差连接(cross-modal residual connections)和互补感知监督(complementarity-aware supervisions)对RGB特征和深度特征进行互补融合,并给出当前尺度下的预测结果。最大尺度下的预测结果则为网络输出的最终预测结果。

[0131] 这种方案直接按照RGB的方式对深度图像进行特征提取,并融合到RGB特征中,由于深度图像具有低分辨率、高噪声、缺乏纹理等特点,因此难以有效地对显著性检测结果进行改进。对于复杂场景,比如低对比度、小物体、多个物体的场景,效果并不显著。一方面,在低对比度的场景,图像的边缘信息缺失或模糊,而图像边缘由于具有分割物体的特性并被大量地应用在显著性物体检测问题上。神经网络在训练过程中无可避免地学习了图像的边缘信息,并以此作为做决策的知识。因此,对于低对比度的场景,模糊的边缘对显著性检测起反作用。另一方面,对于小物体、多个物体等场景,存在多个尺度,而这种方案仅在每个尺度单独做融合,显著性物体检测效果并不理想。

[0132] 而本申请实施例提供的RGBD显著性检测方法、神经网络模型以及相关装置能够实现在诸如低对比度、小物体、多个物体等复杂场景下也能取得较好的显著性检测效果。

[0133] 下面描述本申请实施例中用于RGBD显著性检测的神经网络系统,该神经网络系统

可包括神经网络模型和融合模块,该融合模块可以是本申请实施例提供的流动金字塔模型,可选的,也可以是传统金字塔模型。参见图5,图5是本申请实施例提供的一种神经网络系统的示意性图,如图5所示,该神经网络系统中的神经网络模型可以根据VGGNet (Visual Geometry Group Net)中的VGG-16网络架构为基础构型。VGG-16可对RGB图像中的特征进行提取,经过神经网络的多个阶段可得到不同尺度的特征图。为引入深度图像中的深度信息,本申请实施例设计了特征增强模块(feature-enhanced module, FEM), FEM用于结合RGB信息和深度信息,获得跨模态特征。在不同的网络阶段获得不同尺度的跨模态特征后,还可采用本申请实施例设计的流动金字塔框架(即流动金字塔模型)对多尺度的跨模态特征进一步处理和融合。详细描述如下:

[0134] (1) 该神经网络模型可包括卷积层、池化层、全连接层和特征增强模块(FEM)。具体的,该神经网络模型中含有16个有模型参数的层,示例性地如图5所示,该神经网络模型可包括13个卷积层(有模型参数)、3个全连接层(有模型参数)和5个池化层(无模型参数),示例性地,3*3的小型卷积核和2*2的最大池化层反复堆叠,即,卷积层和池化层形成多种组合。每个卷积层均使用3*3的小型卷积核,每个池化层均为2*2的最大池化层。需要说明的是,上述示例仅用于解释本申请,本申请中的卷积核和池化层还可以是其他的设计规格。

[0135] 为了获得不同尺度的图像特征,可根据“卷积层+池化层”组合将网络划分出5个网络阶段(如图示中的阶段1-阶段5),示例性地,阶段1可包括2个卷积层和1个池化层,阶段2可包括2个卷积层和1个池化层,阶段3可包括3个卷积层和1个池化层,阶段4可包括3个卷积层和1个池化层,阶段5可包括3个卷积层和1个池化层。图5中各阶段中,卷积层和池化层排列顺序仅用作示例,需理解的,本申请中的各阶段的卷积层和池化层还可以是其他的排列顺序,本申请不做限定。

[0136] 本申请中所描述的网络阶段又可称为特征提取模块,也就是说在上述实例中,神经网络模型可包括5个特征提取模块,其中第1个特征提取模块包括2个卷积层和1个池化层;第2个特征提取模块包括2个卷积层和1个池化层;第3个特征提取模块包括3个卷积层和1个池化层;第4个特征提取模块包括3个卷积层和1个池化层;第5个特征提取模块包括3个卷积层和1个池化层。

[0137] RGB图像从神经网络模型的输入层(图5未示)输出到该神经网络模型,RGB图像在神经网络中前向传播,从输入层进一步传输到阶段1的各个层进行依次处理,如在卷积层即进行卷积处理,在池化层即进行池化处理,后续依次经过阶段2-阶段5进行卷积/池化处理。对于不同阶段的“卷积层+池化层”组合,卷积核对图像进行过滤,通过卷积处理,提取某个或某些属性的特征,卷积后输出特征映射图。池化层对该特征映射图进行池化,进一步减少参数的数量(特征图尺寸减半),得到特征向量。通常来说,由于卷积和池化过程会不断地丢失一些非显著的信息,以减少特征映射图的参数和空间尺寸,越深的卷积神经网络会提取越具体明显的特征,越浅的网络提取越浅显的特征。所以,通过不同阶段,可在不同尺度上提取RGB图像中的特征,即通过有效地利用VGG-16能用较低成本实现了不同尺度提取特征,有利于保证小物体以及复杂场景的特征显现。多尺度信息能帮助优化有多个物体和小物体的应用场景的检测效果。

[0138] (2) 在不同的网络阶段之间均设置有FEM,不同阶段产生的RGB特征图(或称RGB的模态信息)被输入到位于该阶段后端的FEM。各个FEM一方面可用于对输入FEM的深度图像进

行处理,提取深度图中的对比度先验知识信息(可简称对比度先验),相当于引入了图像边缘信息,从而获得深度对比度增强图。对比度先验用于表征影响图像的对比度的参数或信息。本申请可预先通过提取深度图中的对比度先验进行模型训练,来增强FEM提取深度图中的对比度先验的能力。另一方面可用于在不同的尺度下将RGB特征和深度特征相结合,利用深度对比度增强图来优化RGB特征图,从而获得不同尺度下的跨模态特征。其中,输入至FEM的RGB特征图来源于FEM前端的网络阶段(如以图6中最左端的FEM为例,该网络阶段为阶段1),FEM输出的跨模态特征一方面作为下一网络阶段的输入数据(如以图6中最左端的FEM为例,该网络阶段为阶段2),另一方面作为独立的输出结果应用到后续的金字塔框架中。

[0139] 具体的,FEM可采用卷积层集合提取深度图像的特征(或称深度图像的模态信息),并对深度图像的特征提取深度图中的对比度先验,然后将它与从RGB特征图中的特征相乘,最后将相乘后的结果与RGB特征图相加,得到跨模态特征(或称跨模态特征图,或称跨模态特征信息)。

[0140] 如图6所示,图6是一种FEM中的数据流程的示意图。FEM采用卷积层集合提取深度图中的特征,利用激活函数在深度图的特征中引入非线性特性的表达,使得经激活函数处理后的特征与原深度图的特征之间具有了非线性映射的关系,有利于本申请中图像边缘信息的引入,该激活函数可以是sigmoid函数。然后,使用本申请实施例提供的对比度增强损失函数对其进行监督,最终得到深度对比度增强图。然后可将深度对比度增强图中的特征与从RGB特征图中的特征相乘,最后将相乘后的结果与RGB特征图相加,得到跨模态特征。由于本申请中不同的FEM输入的RGB特征图的抽象程度不同,所以不同的FEM输出不同尺度的跨模态特征。

[0141] 其中,不同FEM采用的卷积层个数可能是各有差异的,例如,本申请实施例中5个阶段后对应FEM模块的卷积层个数可分别为4,4,6,8和10。

[0142] 其中,在对神经网络模型的训练中,监督的含义是:将经激活函数处理获得的深度图的特征应用到对比度增强损失函数,计算对比度增强损失函数的值,观察对比度增强损失函数的值是否到达局部最小值,如果未达到局部最小,则可更新该FEM的卷积层集合中的各卷积层的卷积核参数,继续梯度下降迭代,直到达到局部最小;达到局部最小时该FEM的卷积核参数即为经训练最终确定的卷积核参数。

[0143] 本申请实施例中采用的对比度增强损失函数如公式(2-1)所示:

$$[0144] \quad l_c = a_1 l_f + a_2 l_b + a_3 l_w \quad (2-1)$$

[0145] 其中, a_1 、 a_2 和 a_3 可分别设置为5、5和1。

[0146] l_f 表示监督前景部分的训练,使前景部分数值分布更加均匀,数学描述如公式(2-2)所示:

$$[0147] \quad l_f = -\log\left(1 - 4 \times \sum_{p \in F} \frac{(p - \widehat{p}_f)^2}{N_f}\right) \quad (2-2)$$

[0148] 其中,F为前景点集合, \widehat{p}_f 为前景点均值, N_f 为前景点个数。

[0149] l_b 表示监督背景部分的训练,使背景部分数值分布更加均匀,数学描述如公式(2-3)所示:

$$[0150] \quad l_b = -\log\left(1 - 4 \times \sum_{p \in B} \frac{(p - \widehat{p}_b)^2}{N_b}\right) \quad (2-3)$$

[0151] 其中, B 为背景点集合, \widehat{p}_b 为背景点均值, N_b 为背景点个数。

[0152] l_w 表示监督全图训练, 使前景背景均值差变大, 对比度增强, 数学描述如公式 (2-4) 所示:

$$[0153] \quad l_w = -\log(\widehat{p}_f - \widehat{p}_b)^2 \quad (2-4)$$

[0154] 其中, \widehat{p}_f 为前景点均值, \widehat{p}_b 为背景点均值, 数学描述见公式 (2-5) 和公式 (2-6)。

$$[0155] \quad \widehat{p}_f = \sum_{p \in F} \frac{p}{N_f} \quad (2-5)$$

$$[0156] \quad \widehat{p}_b = \sum_{p \in B} \frac{p}{N_b} \quad (2-6)$$

[0157] 使用本申请实施例提供的对比度增强损失函数对深度特征进行监督, 能够提高深度图的对比度, 引入了图像边缘信息, 获得深度对比度增强图。将该深度对比度增强图与从 RGB 特征图中的特征相乘, 最后将相乘后的结果与 RGB 特征图相加, 可得到跨模态特征。

[0158] 可以看到, 本申请实施例通过引入 FEM 模块提取深度图中的对比度先验, 提高深度图的对比度, 引入了图像边缘信息, 并据此优化 RGB 分支的特征, 使得算法在低对比度场景有较好的性能, 对深度图的低分辨率、噪声鲁棒性高。通过考虑不同尺度之间的 RGB 特征和深度特征的互补, 有效地利用了多尺度信息以及对比度先验信息, 从而能帮助优化低对比度或构图复杂应用场景的检测效果。

[0159] (3) 在不同的网络阶段获得不同尺度的跨模态特征后, 这些跨模态特征可通过本申请实施例设计的流动金字塔框架进行融合处理。

[0160] 如图 5 所示, 流动金字塔框架的设计描述如下: 首先搭建出多层 (例如, 5 层) 的金字塔框架, 金字塔的每层的图像集包括一个或多个尺度特征 (一个尺度特征可认为是一个图像), 且各层的尺度特征的数量依次递减 (即从图示中的倒金字塔底到顶逐层递减), 例如金字塔的第 1 层 (金字塔的底层) 包括 5 个尺度特征 (即对应 5 个阶段后的 FEM 分别输出的跨模态特征), 第 2 层包括 4 个尺度特征, 第 3 层包括 3 个尺度特征, 第 4 层包括 2 个尺度特征, 第 5 层 (金字塔的顶层) 包括 1 个尺度特征。其中第 1 层的 5 个尺度特征分别对应 5 个特征提取模块输出的 5 个跨模态特征图, 第 5 层的 1 个尺度特征即为所述金字塔最终输出的融合结果。

[0161] 向金字塔引入流动连接, 即对于金字塔的第 2-5 层, 按照从第 2 层到第 5 层的顺序, 每层的图像集中的每个尺度特征都由前一更低层的图像集中尺度不大于该尺度特征的尺度的所有尺度特征融合而成。不同特征融合的方式为像素级别的相加 (例如图示中箭头的汇聚为相加)。

[0162] 如图示中, 第 2 层的第一个尺度特征 (L2-1a) 为根据“阶段 1-阶段 5”后分别对应的 5 个 FEM 所输出的跨模态特征融合而成。第 2 层的第二个尺度特征 (L2-2a) 为根据“阶段 2-阶段 5”后分别对应的 4 个 FEM 所输出的跨模态特征融合而成。第 2 层的第三个尺度特征 (L2-3a) 为根据“阶段 3-阶段 5”后分别对应的 3 个 FEM 所输出的跨模态特征融合而成。第 2 层的第四个尺度特征 (L2-4a) 为根据“阶段 4-阶段 5”后分别对应的 2 个 FEM 所输出的跨模态特征融合而成。其他金字塔层的尺度特征的融合方式以此类推。

[0163] 通过引入流动连接, 而使多尺度的跨模态特征得到更为充分的互补和融合, 通过逐层融合后的特征最后被输出 (例如被第 5 层的金字塔层融合处理后输出), 输出结果被输入到后续的神经网络层进行后续处理。

[0164] 需要说明的是,上述以VGG-16为基础构型的神经网络模型作用作示例而非限定,基于本申请实施例的技术思想,其他的神经网络变形结构同样隶属于本申请的保护范围。

[0165] 还需要说明的是,图5所示的流动金字塔以5层为示例进行方案的解释,在其他可能的实施例中,流动金字塔还可以是其他的层数,这里不做限定。

[0166] 可以看到,本申请实施例提供的流动金字塔引入了更多的跨尺度连接来获得特征融合结果,使得多尺度的跨模态特征融合与互补更为充分。对于任一尺度特征而言,其拥有了更多尺度的RGB特征的信息,不同尺度之间的跨模态特征融合,促进了RGB特征与深度特征更好地互补。所以本申请实施例引入流动金字塔能够保证其RGB-D图像数据的兼容性和互补性,进一步帮助优化低对比度或构图复杂应用场景的检测效果,使得RGB-D显著性检测效果更加具有普适性。

[0167] (4) 经金字塔融合处理的数据可被输入到全连接层,本申请实施例中的全连接层的数量可以是3个,各个全连接层的神经元数量例如可以为4096、4096和1000。通过全连接层对数据进行整合和降维,从而得到与显著性分类总数相同的归一化的检测结果图。最后将检测结果图输出,从而实现了对RGB-D图像数据的显著性检测。

[0168] 在实际应用中,对检测结果图后续还可以应用于图像分割,目标识别,视觉跟踪,图像检索,图像编辑,图像语义分割,虚拟大光圈拍照,翻译,文本分析等等应用场景中,这里不做限定。

[0169] 参见图7,图7是本申请实施例提供的又一种神经网络模型的示意性图,图7所示的神经网络模型同样可以根据VGGNet中的VGG-16网络架构为基础构型,该神经网络模型包括5个网络阶段,不同阶段之间采用FEM连接,关于5个网络阶段以及FEM的相关内容可参考图5实施例的描述,这里不再赘述。各个FEM输出的不同尺度的跨模态特征采用金字塔进行融合处理。图7所示的神经网络模型与图5所示的神经网络模型的区别在于:图5采用本申请设计的流动金字塔框架来对不同尺度的跨模态特征进行融合处理,而图7采用传统的金字塔框架来对不同尺度的跨模态特征进行融合处理。如图7所示,在传统的金字塔框架中,即对于金字塔的第2-4层,每层的每个尺度特征都由前一层的图像集中同尺度的尺度特征和尺度小一级的尺度特征两者融合而得,特征融合的方式为像素级别的相加(例如图示中箭头的汇聚为相加)。

[0170] 如图示中,第2层的第一个尺度特征(L2-1b)为根据“阶段1和阶段2”后分别对应的2个FEM所输出的跨模态特征融合而成。第2层的第二个尺度特征(L2-2b)为根据“阶段2和阶段3”后分别对应的2个FEM所输出的跨模态特征融合而成。第2层的第三个尺度特征(L2-3b)为根据“阶段3和阶段4”后分别对应的2个FEM所输出的跨模态特征融合而成。第2层的第四个尺度特征(L2-4b)为根据“阶段4和阶段5”后分别对应的2个FEM所输出的跨模态特征融合而成。其他金字塔层的尺度特征的融合方式以此类推。

[0171] 通过传统金字塔使得多尺度的跨模态特征进行融合,在一定程度上也能保证输出结果的兼容性和互补性。

[0172] 下面以图8的测试结果为例对不同设计情况的显著性检测效果进行不对比说明。图8中的左侧两列图分别为RGB图像和对应的深度图像。右侧五列图中:

[0173] (B) 列图:表示本申请图5实施例的神经网络模型去掉FEM模块和流动金字塔框架后,只对RGB图像进行显著性检测获得的检测结果图;

[0174] (B+C) 列图:表示本申请图5实施例的神经网络模型去掉流动金字塔框架后,借助对比度先验进行显著性检测获得的检测结果图;

[0175] (B+C+P) 列图:表示通过本申请图7实施例的神经网络模型(即包含FEM和传统金字塔框架)进行显著性检测获得的检测结果图;

[0176] (B+C+FP) 列图:表示通过本申请图5实施例的神经网络模型(即包含FEM和流动金字塔框架)进行显著性检测获得的检测结果图;

[0177] (GT) 列图:表示通过人工标注的显著目标结果图。

[0178] 通过对比可以看出,本申请实施例通过深度对比度先验,提高了深度图的对比度,得到更为清晰的目标边缘。通过多尺度跨模态特征融合,更好地实现了RGB和深度的跨模态互补。图5实施例中结合FEM和流动金字塔的方案由于同时有效利用深度对比度先验和多尺度的跨模态特征,更为准确地检测显著性物体,在图8所示的各种对比方案中效果最好。图7实施例中结合FEM和传统金字塔的方案也在一定程度上满足显著性检测的需求,但是效果不如图5的方案出色。

[0179] 本申请各实施例描述的神经网络模型在实际应用之前,通常会基于大量的样本数据进行自动深度学习,采用相关训练算法和反向传播算法进行大量训练,从而不断地调整神经网络模型中的参数设定,提高RGBD显著性检测的性能。下面以图5实施例描述的神经网络模型为例描述模型训练方法。

[0180] 本申请实施例中,神经网络模型可通过监督学习(Supervised Learning)的方法进行训练,在需要进行训练时,对数据库输入大量的作为样本数据的RGB-D图像数据,样本数据中可通过人工标注的方式描绘显著性目标的轮廓。神经网络模型分别获取样本数据的RGB图像和深度图像进行训练,在每个训练周期,基于样本数据的人工标注轮廓和预测结果图,计算损失函数值。神经网络模型可通过误差反向传播(back propagation, BP)算法从全连接层由后往前传递梯度到各网络阶段(如阶段5-阶段1)的卷积层以及各FEM,进而修正初始的神经网络模型中相关层的参数的大小,使得神经网络模型的重建误差损失越来越小。具体地,前向传递输入信号直至输出会产生误差损失,通过反向传播误差损失信息来更新初始的神经网络模型中的模型参数,从而使误差损失收敛。反向传播算法是以误差损失为主导的反向传播运动,旨在得到最优的神经网络模型的参数。

[0181] 基于大量的样本数据进行训练后,能够不断自动调整神经网络模型的一些参数设定,从而使得对RGBD显著性检测结果不断地趋向准确。这些可以通过大量学习自动确定和修正的模型参数例如包括:各网络阶段(如阶段5-阶段1)的卷积层的卷积核(卷积矩阵)、全连接层中的参数,各FEM中的卷积层的卷积核(卷积矩阵),各FEM中的对比度增强损失函数中的参数,等等。

[0182] 经过大量样本数据的模型训练获得神经网络模型后,基于该神经网络模型,下面描述本申请实施例提供的RGBD显著性检测方法,参见图13,该方法包括但不限于如下步骤:

[0183] S101:获取待检测的RGB-D图像数据。RGB-D图像数据可包括RGB图像和对应的深度图像。

[0184] S102:通过N个特征提取模块以及N个特征增强模块(FEM),对RGB图像进行多种不同抽象尺度的特征提取,输出对应N个特征提取模块的N种抽象程度的处理结果。N为大于等于2的整数。

[0185] 具体的,可通过不同的卷积层和池化层的组合实现对RGB图像不同程度的抽象,也就是说,位于神经网络较浅层次的网络层对图像抽象程度低,图像中的细粒度特征得以保留;位于神经网络较深层次的网络层对图像抽象程度高,图像中的粗粒度特征得以保留。通过多种不同抽象程度的特征提取,可得到不同抽象程度、不同尺度的RGB特征。

[0186] 本申请实施例中,一个特征提取模块可为“M层卷积层+K层池化层”的组合,特征提取模块和FEM交替设置,即不同特征提取模块之间通过FEM连接。其中,M为大于等于1的整数,K为大于等于1的整数。对于不同的特征提取模块,M可能是相同的,也可能是各有差异的;N可能是相同的,也可能是各有差异的。

[0187] 示例性地,该神经网络模型为如图5所描述的神经网络模型,可根据“卷积层+池化层”组合将神经网络划分出5个网络阶段(即5个特征提取模块),示例性地,阶段1可包括2个卷积层和1个池化层,阶段2可包括2个卷积层和1个池化层,阶段3可包括3个卷积层和1个池化层,阶段4可包括3个卷积层和1个池化层,阶段5可包括3个卷积层和1个池化层。

[0188] 对于任一特征提取模块,例如N个特征提取模块中的第j个特征提取模块,其对数据的处理可描述如下:第j个特征提取模块对所述第j个特征提取模块的输入数据进行特征提取,具体的,通过M层卷积层的卷积处理和K层池化层的池化处理,获得第j个特征提取模块的处理结果,并将该处理结果输出到N个特征强化模块中的第j个特征强化模块;其中,当j等于1时,输入数据例如为RGB图像(这里的RGB图像例如是经神经网络模型的输入层处理后再输到第1个特征提取模块的),第j个特征提取模块的处理结果即为对该RGB图像进行特征提取所得到的结果;当j不等于1时,输入数据为第j-1个特征强化模块输出的跨模态特征图,第j个特征提取模块的处理结果即为对该跨模态特征图进行特征提取所得到的结果;其中,j为大于等于1且小于等于N的整数;

[0189] 示例性地,对于5个组合,即图5所示的阶段1-阶段5,可分别输出不同抽象尺度RGB特征数据(即RGB图像或跨模态特征图)到各组合后端的FEM。

[0190] S103:分别通过N个特征增强模块中的各个特征增强模块(FEM)提取深度图像的特征,实现对图像对比度的增强,获得各个特征增强模块对应的深度对比度增强图。

[0191] N个特征强化模块中的第j个特征强化模块对数据的处理可描述如下:第j个特征强化模块对导入至所述第j个特征强化模块的深度图像(这里的深度图像例如是经神经网络模型的输入层处理后再输到第j个特征强化模块的)进行特征提取和经激活函数处理,获得深度对比度增强图;然后利用深度对比度增强图优化第j个特征提取模块的处理结果,获得第j个特征强化模块的跨模态特征图并输出。

[0192] 具体的,FEM可采用卷积层集合提取深度图中的特征,卷积层集合包括H层的卷积层,H为大于等于1的整数;不同FEM采用的卷积层个数可能是各有差异的。可利用激活函数在深度图的特征中引入非线性特性的表达,该激活函数可以是sigmoid函数。卷积层集合的各卷积层的卷积核参数(如权重矩阵)可以是在对神经网络模型进行模型训练过程中利用本申请实施例中提供的对比度增强损失函数对其进行监督而得到的。经过上述处理过程可实现利用对比度先验提高深度图的对比度,使得图像边缘信息更为显著,最终得到深度对比度增强图。

[0193] 具体的,对第j个特征强化模块进行监督以确定卷积层集合的各个卷积层的卷积核参数的过程可描述如下:在模型训练过程中,可将导入至第j个特征强化模块的深度图像

通过第j个特征强化模块的卷积层集合的各个卷积层进行卷积处理,以及通过激活函数处理,获得处理后的深度特征图;然后利用对比度增强损失函数对该深度特征图进行运算,得到对比度增强损失函数的值;在对比度增强损失函数的值未达到局部最小值的情况下,更新卷积层集合的各个卷积层的卷积核参数,继续梯度下降迭代,直到达到局部最小值。在达到局部最小值对应的卷积核参数即作为卷积层集合的各个卷积层的卷积核参数,以应用到后续的应用场景(如显著性检测)中。

[0194] 其中,对比度增强损失函数的相关内容可参考上文公式(2-1)至公式(2-6)的相关描述,这里不再赘述。

[0195] S104:将输入各个特征强化模块的不同抽象尺度的处理结果与各个特征强化模块的深度对比度增强图相乘,并将相乘后的结果与对应输入各个特征强化模块的处理结果相加,获得对应特征强化模块的不同抽象尺度的跨模态特征。

[0196] 也即是说,对于第j个特征增强模块,将第j个特征增强模块处理获得的深度对比度增强图与输入到第j个特征增强模块的处理结果(即第j个特征提取模块输出的处理结果)相乘,获得相乘后的结果;将相乘后的结果与上述处理结果相加,获得第j个特征强化模块的跨模态特征图。

[0197] 需要说明的是,S103和S104的具体实现可参考图5实施例中关于FEM的相关描述,为了说明书的简洁,这里不再赘述。

[0198] S105:对不同抽象尺度的跨模态特征采用金字塔模型进行特征融合,获得融合结果。

[0199] 其中,该金字塔模型包括具有多个层的金字塔,各层对应有自己的图像集,顶层的图像集包括一个图像。底层的图像集包括N个图像,N个图像与N个特征强化模块输出的N个跨模态特征图一一对应。从金字塔的底层到金字塔的顶层的各个层的图像数量依次递减。

[0200] 在一些具体实施例中,该金字塔模型可为本申请实施例提供的第一种金字塔模型,所述第一种金字塔模型又可称为流动金字塔模型,在利用流动金字塔模型进行数据融合过程中,按照从所述金字塔的底层到所述金字塔的顶层的顺序,获得除底层外的每一层的图像集;其中,除底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度不大于所述任一图像的尺度的所有图像融合而成,不同图像融合的方式具体为像素级别的相加。

[0201] 这样,各个尺度的跨模态特征进入流动金字塔后,逐层相互作用至金字塔的顶层,实现了充分的互补和融合,金字塔的顶层的融合图像即为最终得到融合结果。

[0202] 在又一些具体实施例中,该金字塔模型也可以是本申请实施例提供的第二种金字塔模型,在利用第二种金字塔模型进行数据融合过程中,按照从所述金字塔的底层到所述金字塔的顶层的顺序,获得除底层外的每一层的图像集;其中,除底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度同所述任一图像相同的一个图像和尺度比所述任一图像小的一个图像两者融合而成,不同图像融合的方式具体为像素级别的相加。

[0203] 第二种金字塔模型例如可以是传统的金字塔模型,这种情况下,除底层外的任一层的图像集中,任一图像为根据前一更低层的图像集中尺度同所述任一图像相同的一个图像和尺度比所述任一图像小一级的图像两者融合而成。

[0204] 各个尺度的跨模态特征进入第二种金字塔模型后,逐层相互作用至金字塔的顶

层,也能在一定程度上实现了互补和融合,金字塔的顶层的融合图像即为最终得到融合结果。

[0205] 需要说明的是,S105的具体实现可参考图5实施例中关于流动金字塔框架的相关描述,或参考图7实施例中关于传统金字塔框架的相关描述,为了说明书的简洁,这里不再赘述。

[0206] S106:根据融合结果获得显著性检测的检测结果图。

[0207] 具体的,可将融合结果输入到全连接层,通过全连接层对数据进行整合和降维,从而获得检测结果图。最后将检测结果图输出,从而实现了RGB-D图像数据的显著性检测。

[0208] 需要说明的是,S106的具体实现可参考图5实施例中关于全连接层的相关描述,为了说明书的简洁,这里不再赘述。

[0209] 可以看到,本申请实施例通过实现了不同尺度提取特征,有利于保证小物体以及复杂场景的特征显现。通过FEM增加深度图的对比度先验,提高了深度图的对比度,得到更为清晰的目标边缘,能够优化低对比度或构图复杂应用场景的检测效果。通过多尺度的跨模态特征融合,更好地实现了RGB信息和深度信息的跨模态互补,保证其RGB-D图像数据的兼容性和互补性。所以,实施本申请实施例能够实现更为准确地检测显著性物体,有利于提高显著性检测的普适性,在诸如低对比度、小物体、多个物体等复杂场景下也能取得较好的显著性检测效果。

[0210] 进一步地,可以参见图14,图14展示了本申请提供的RGBD显著性检测方法 with 近几年其他RGBD显著性检测方法的技术效果对比图。图14对简单场景、低对比度场景、构图复杂场景、小物体场景和多个物体场景下对RGB-D图像(包括RGB图和对应的深度图)的检测结果图做了分类呈现。可以看到,由于本申请的方法合理利用了深度对比度先验和多尺度跨模态信息,使得本申请的方法在上述各种场景下都能取得了优异的检测结果,其检测效果皆优于其他RGBD显著性检测方法(如图示中的PCF、CTMF、MDSF、DF、CDCP、SE、LBE等方法)。

[0211] 上述详细阐述了本申请实施例的方法,下面提供了本申请实施例的相关装置。

[0212] 参见图9,图9是本申请实施例提供的一种芯片硬件结构图。如图9所示,神经网络处理器(NPU)50作为协处理器挂载到主CPU(Host CPU)上,由Host CPU分配任务。NPU的核心部分为运算电路50,通过控制器504控制运算电路503提取存储器中的矩阵数据并进行乘法运算。

[0213] 在一些实现中,运算电路503内部包括多个处理单元(Process Engine,PE)。在一些实现中,运算电路503是二维脉动阵列。运算电路503还可以是一维脉动阵列或者能够执行例如乘法和加法这样的数学运算的其它电子线路。在一些实现中,运算电路503是通用的矩阵处理器。

[0214] 举例来说,假设有输入矩阵A,权重矩阵B,输出矩阵C。运算电路从权重存储器502中取矩阵B相应的数据,并缓存在运算电路中每一个PE上。运算电路从输入存储器501中取矩阵A数据与矩阵B进行矩阵运算,得到的矩阵的部分结果或最终结果,保存在累加器(accumulator)508中。

[0215] 统一存储器506用于存放输入数据(例如,RGB-D图像数据)以及输出数据(例如,检测结果图)。权重数据直接通过存储单元访问控制器(Direct Memory Access Controller)505,DMAC被搬运到权重存储器502中。输入数据也通过DMAC被搬运到统一存储

器506中。

[0216] 总线接口单元 (Bus Interface Unit, BIU) 510, 用于AXI总线与DMAC和取指存储器 (Instruction Fetch Buffer) 509的交互。

[0217] 总线接口单元510具体用于取指存储器509从外部存储器获取指令, 还用于存储单元访问控制器505从外部存储器获取输入矩阵A或者权重矩阵B的原数据。

[0218] DMAC主要用于将外部存储器DDR中的输入数据搬运到统一存储器506或将权重数据搬运到权重存储器502中或将输入数据数据搬运到输入存储器501中。

[0219] 向量计算单元507多个运算处理单元, 在需要的情况下, 对运算电路的输出做进一步处理, 如向量乘, 向量加, 指数运算, 对数运算, 大小比较等等。主要用于神经网络中非卷积/FC层网络计算, 如池化 (Pooling), 批归一化 (Batch Normalization), 局部响应归一化 (Local Response Normalization) 等。

[0220] 在一些实现中, 向量计算单元507将经处理的输出的向量存储到统一缓存器506。例如, 向量计算单元507可以将非线性函数应用到运算电路503的输出, 例如累加值的向量, 用以生成激活值。在一些实现中, 向量计算单元507生成归一化的值、合并值, 或二者均有。在一些实现中, 处理过的输出的向量能够用作到运算电路503的激活输入, 例如用于在神经网络中的后续层中的使用。

[0221] 控制器504连接的取指存储器 (instruction fetch buffer) 509, 用于存储控制器504使用的指令;

[0222] 统一存储器506, 输入存储器501, 权重存储器502以及取指存储器509均为On-Chip存储器。外部存储器私有于该NPU硬件架构。

[0223] 其中, 图5和图7所示的神经网络模型 (神经网络系统) 中各层的运算可以由向量计算单元507执行。

[0224] NPU50的各部件具体可用于实现图13中所描述的方法, 为了说明书的简洁, 这里不再赘述。

[0225] 基于相同的申请构思, 本申请实施例还提供了一种用于RGBD显著性检测的装置600, 参见图10, 该装置600包括数据神经网络模型601和融合模块604, 神经网络模型601进一步包括N个交替设置的特征提取模块602和特征增强模块 (FEM模块) 603, N为大于等于2的整数。各个模块的功能可描述如下:

[0226] 所述神经网络模型用于, 对获取的RGB图像进行处理, 以及对与所述RGB图像配准的深度图像进行处理, 获得所述N个特征强化模块的N个跨模态特征图; N为大于等于2的整数;

[0227] 其中, 所述N个特征提取模块中的第j个特征提取模块具体用于: 对所述第j个特征提取模块的输入数据进行特征提取, 获得所述第j个特征提取模块的处理结果并输出到所述N个特征强化模块中的第j个特征强化模块; 其中, 当j等于1时, 所述输入数据为所述RGB图像, 当j不等于1时, 所述输入数据为第j-1个特征强化模块输出的跨模态特征图; j为大于等于1且小于等于N的整数;

[0228] 其中, 所述N个特征强化模块中的第j个特征强化模块具体用于: 对导入至所述第j个特征强化模块的深度图像进行特征提取, 获得深度对比度增强图; 利用所述深度对比度增强图优化所述第j个特征提取模块的处理结果, 获得所述第j个特征强化模块的跨模态特

征图；

[0229] 所述融合模块用于,对所述N个跨模态特征图进行融合,获得显著性检测的检测结果图。

[0230] 装置600的各功能模块具体可用于实现图13中所描述的方法,为了说明书的简洁,这里不再赘述。

[0231] 参见图11,图11为本申请实施例提供的一种计算节点700与用户及运维人员交互的示意图。计算节点700可以包括处理器710以及存储器720。其中,处理器用于运行数据神经网络模型601和融合模块604。存储器720用于存储RGB-D图像数据等等。计算节点700还提供了两种对外的接口界面,分别是面向语义描述系统的维护人员的管理界面740以及面向用户的用户界面750。其中,接口界面的形态可以是多样的,例如web界面、命令行工具、REST接口等。

[0232] 在本申请具体的实施例中,管理界面用于供维护人员可以通过输入大量样本数据(例如有标注的RGB-D图像数据),以用于对神经网络模型进行训练。

[0233] 在本申请具体的实施例中,用户界面用于供用户输入需要被进行显著性检测的RGB-D图像数据,并且,通过用户界面向用户输出检测结果图。

[0234] 应当理解,计算节点700仅为本申请实施例提供的一个例子,并且,计算节点700可具有比示出的部件更多或更少的部件,可以组合两个或更多个部件,或者可具有部件的不同配置实现。

[0235] 参见图12,图12是本申请实施例提供的另一计算节点800的结构示意图。计算节点800包括一个或多个处理器811、通信接口812和存储器813。其中,处理器811、通信接口812和存储器813之间可以通过总线824连接。

[0236] 处理器811包括一个或者多个通用处理器,其中,通用处理器可以是能够处理电子指令的任何类型的设备,包括中央处理器(Central Processing Unit,CPU)、微处理器、微控制器、主处理器、控制器以及ASIC(Application Specific Integrated Circuit,专用集成电路)等等。处理器811执行各种类型的数字存储指令,例如存储在存储器813中的软件或者固件程序,它能使计算节点800提供较宽的多种服务。例如,处理器811能够执行程序或者处理数据,以执行本文讨论的方法的至少一部分。处理器811中可以运行数据神经网络模型601和融合模块604。

[0237] 通信接口812可以为有线接口(例如以太网接口),用于与其他计算节点或用户进行通信。

[0238] 存储器813可以包括易失性存储器(Volatile Memory),例如随机存取存储器(Random Access Memory,RAM);存储器也可以包括非易失性存储器(Non-Volatile Memory),例如只读存储器(Read-Only Memory,ROM)、闪存存储器(Flash Memory)、硬盘(Hard Disk Drive,HDD)或固态硬盘(Solid-State Drive,SSD)存储器还可以包括上述种类的存储器的组合。存储器813可以存储有程序代码以及数据(例如RGB-D图像数据、检测结果图)。其中,程序代码例如包括各模块的代码。

[0239] 其中,所述处理器811通过调用存储器813中的程序代码,用于执行以下步骤:

[0240] 获取RGB图像和与所述RGB图像配准的深度图像;

[0241] 将所述RGB图像导入到包含N个特征提取模块和N个特征强化模块的神经网络模型

进行处理,以及将所述深度图像导入所述神经网络模型进行处理,获得所述N个特征强化模块的N个跨模态特征图;N为大于等于2的整数;

[0242] 其中,所述N个特征提取模块中的第j个特征提取模块的处理包括:对所述第j个特征提取模块的输入数据进行特征提取,获得所述第j个特征提取模块的处理结果并输出到所述N个特征强化模块中的第j个特征强化模块;其中,当j等于1时,所述输入数据为所述RGB图像,当j不等于1时,所述输入数据为第j-1个特征强化模块输出的跨模态特征图;j为大于等于1且小于等于N的整数;

[0243] 其中,所述N个特征强化模块中的第j个特征强化模块的处理包括:对导入至所述第j个特征强化模块的深度图像进行特征提取,获得深度对比度增强图;利用所述深度对比度增强图优化所述第j个特征提取模块的处理结果,获得所述第j个特征强化模块的跨模态特征图;

[0244] 对所述N个跨模态特征图进行融合,获得检测结果图。

[0245] 处理器811具体可用于执行图13中所描述的方法,为了说明书的简洁,这里不再赘述。

[0246] 可以看到,本申请实施例中的上述各装置通过实现不同尺度提取特征,有利于保证小物体以及复杂场景的特征显现。通过FEM提取深度图的对比度先验,提高了深度图的对比度,得到更为清晰的目标边缘,能够优化低对比度或构图复杂应用场景的检测效果。通过多尺度的跨模态特征融合,更好地实现了RGB信息和深度信息的跨模态互补,保证其RGB-D图像数据的兼容性和互补性。所以,实施本申请实施例能够实现更为准确地检测显著性物体,有利于提高显著性检测的普适性,在诸如低对比度、小物体、多个物体等复杂场景下也能取得较好的显著性检测效果。

[0247] 在上述实施例中,可以全部或部分地通过软件、硬件、固件或者任意组合来实现。当使用软件实现时,可以全部或者部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令,在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络或其他可编程装置。所述计算机指令可存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,所述计算机指令可以从一个网络站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线)或无线(例如红外、微波等)方式向另一个网络站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质,也可以是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质(例如软盘、硬盘、磁带等)、光介质(例如DVD等)、或者半导体介质(例如固态硬盘)等等。

[0248] 在上述实施例中,对各个实施例的描述各有侧重,某个实施例中沒有详述的部分,可以参见其他实施例的相关描述。

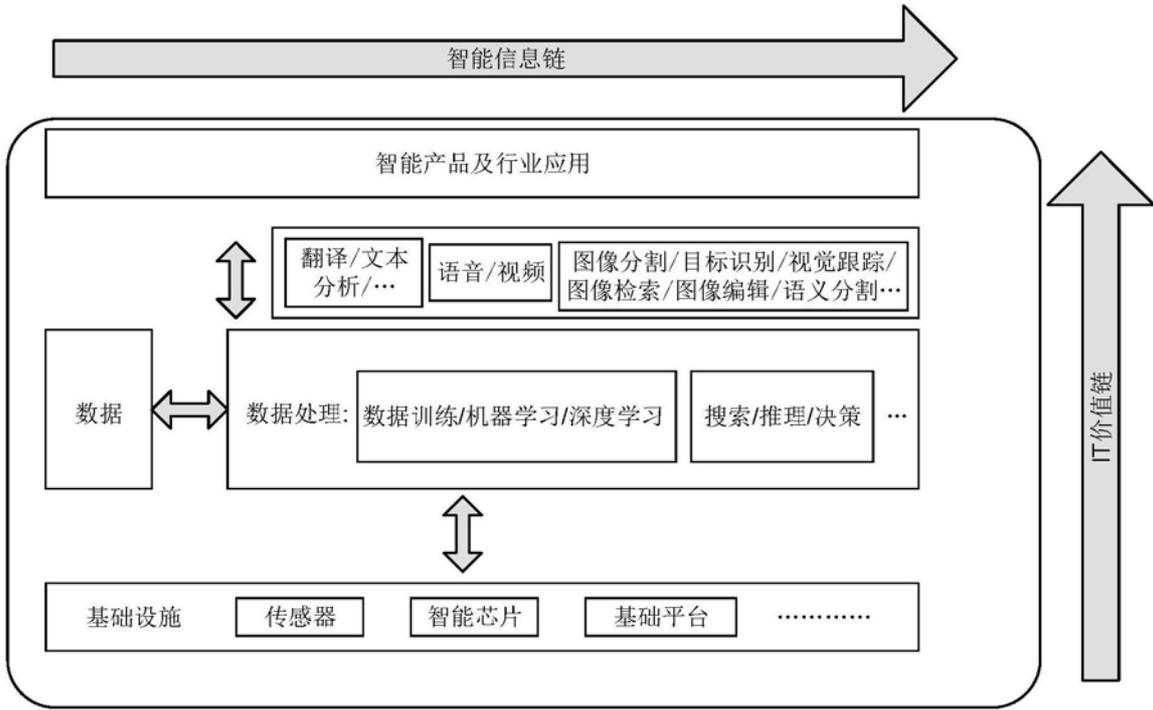


图1

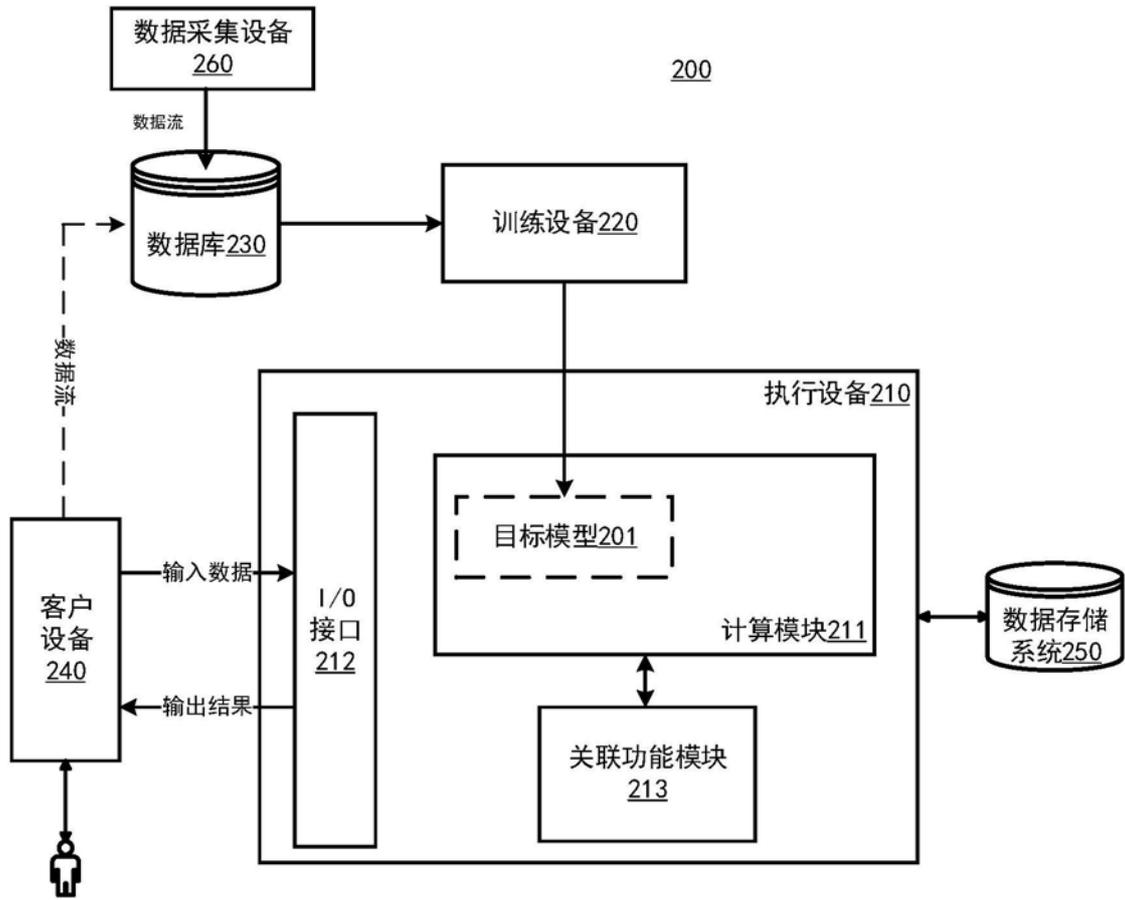


图2a

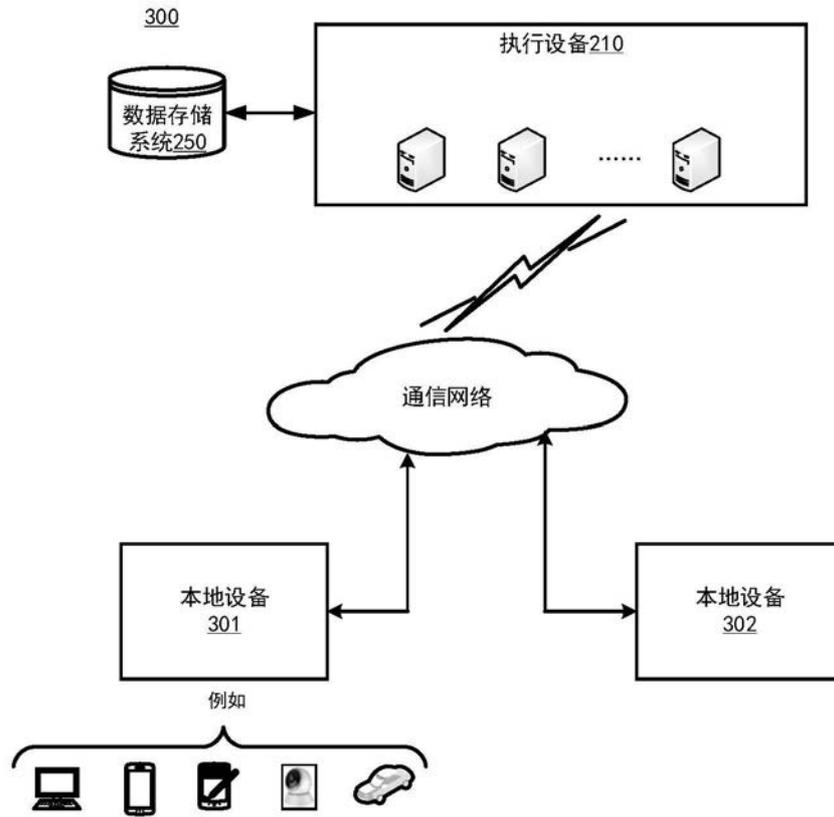


图2b

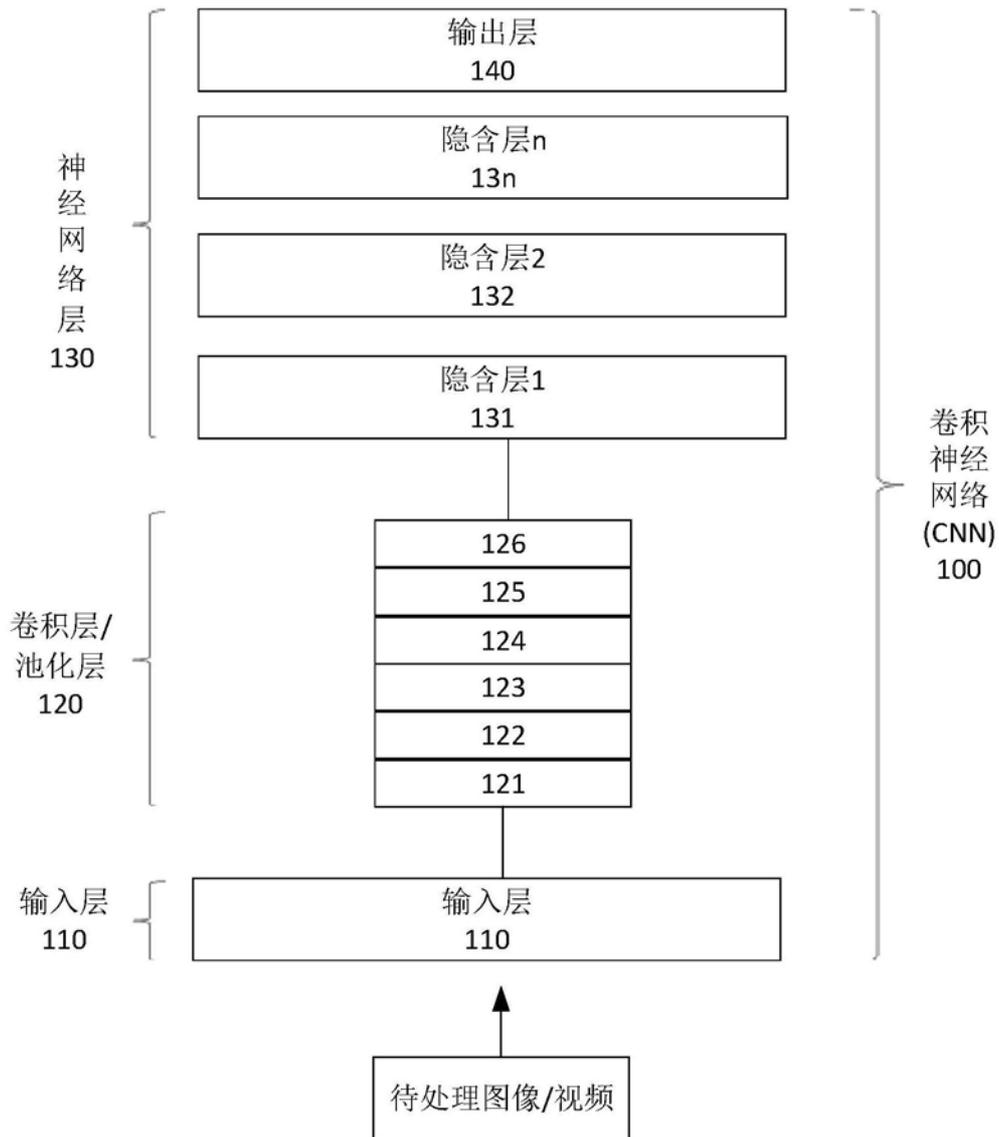


图3

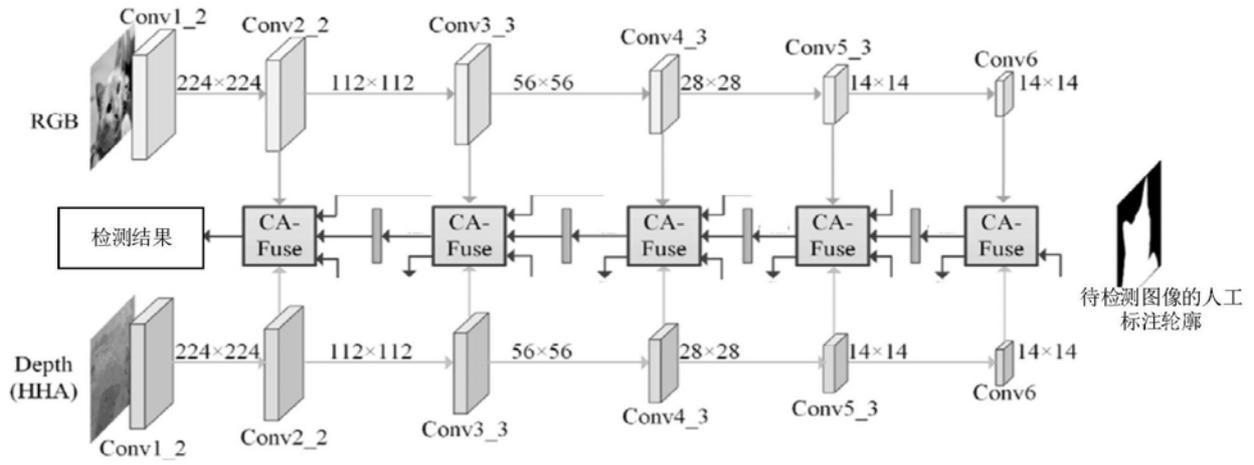


图4

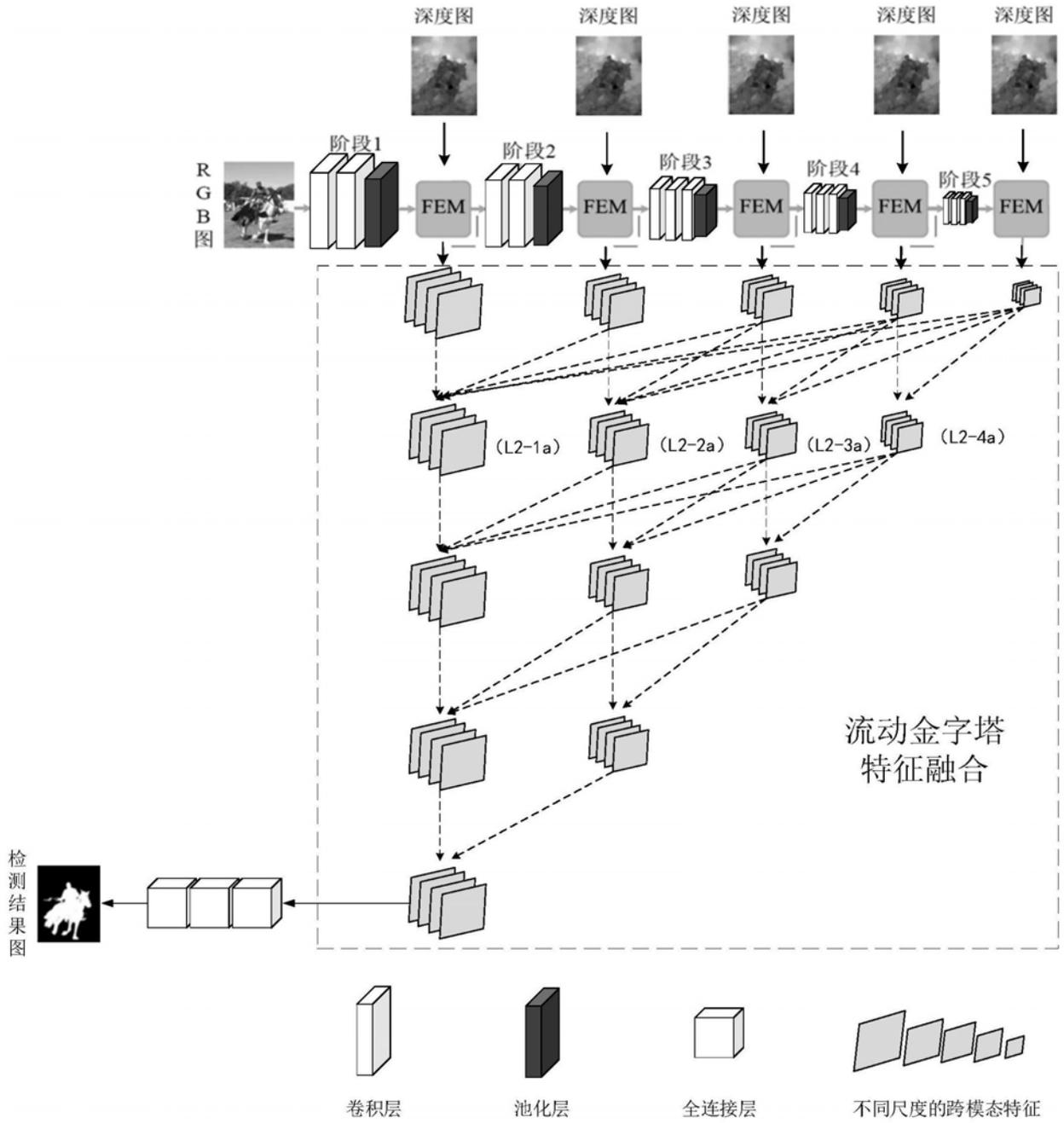


图5

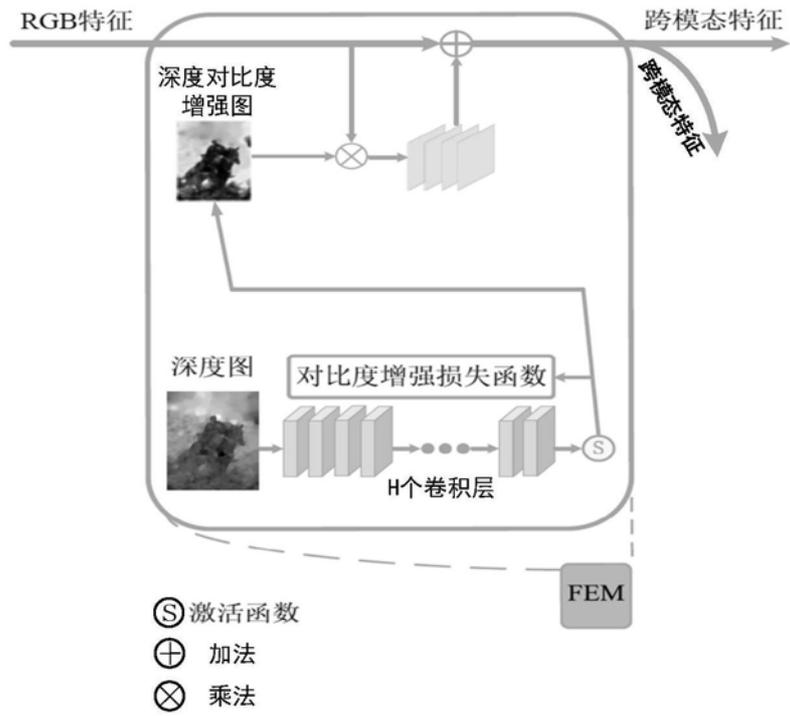


图6

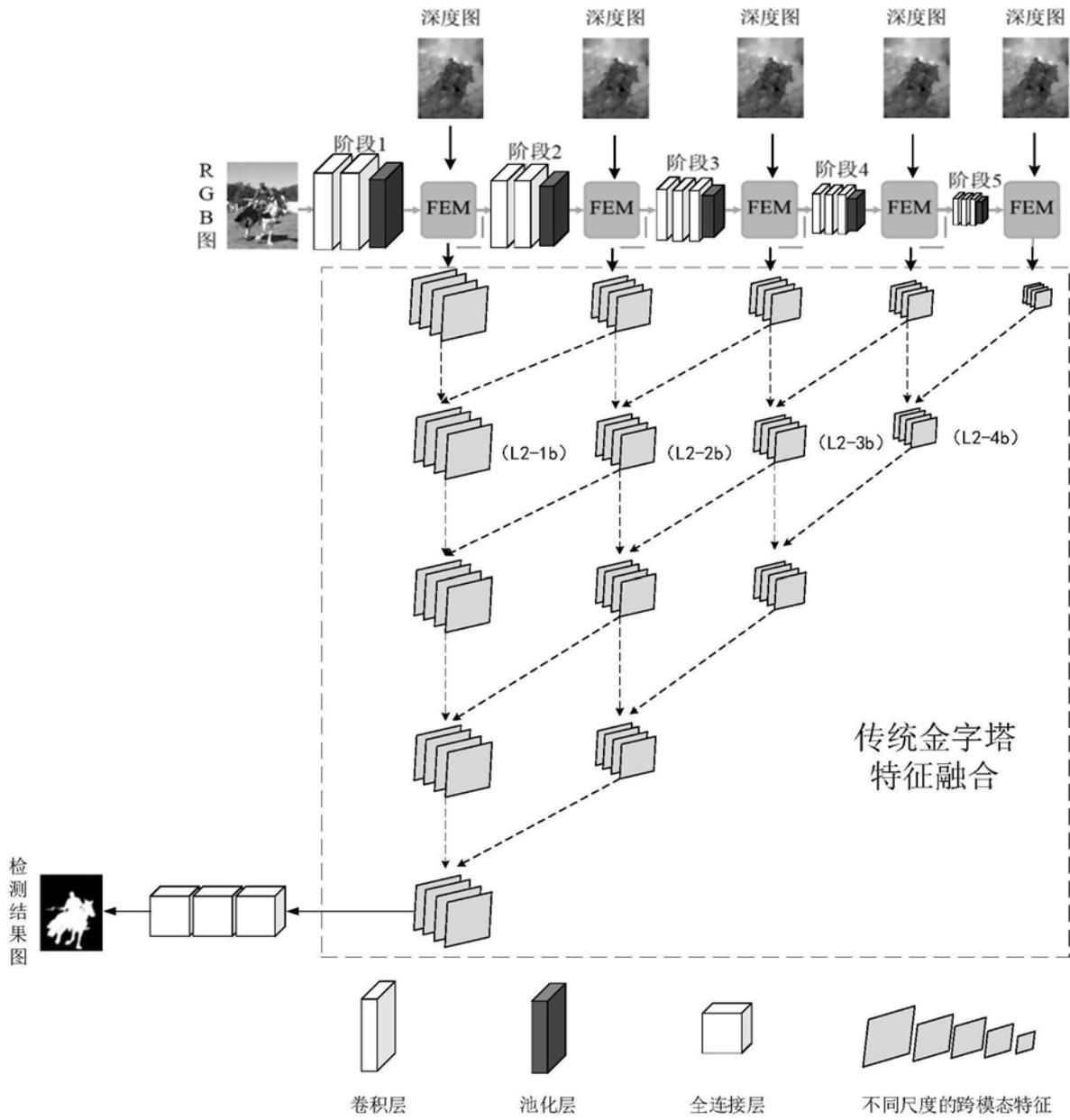


图7

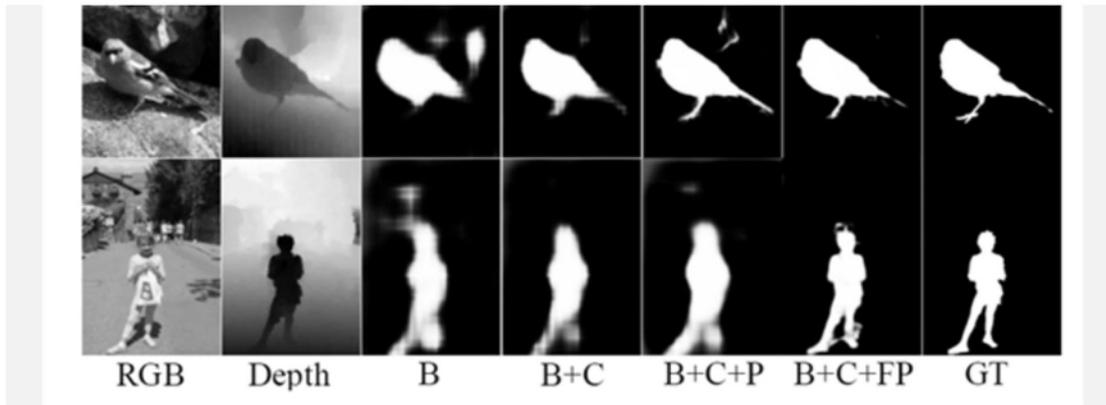


图8

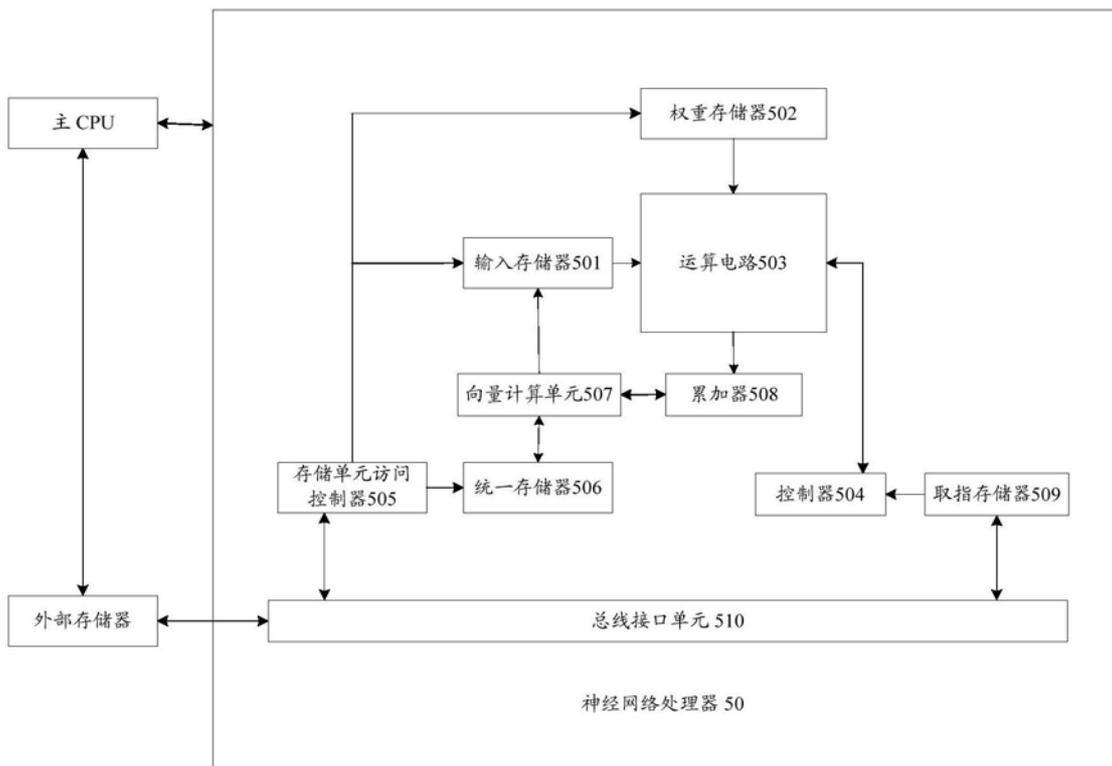


图9

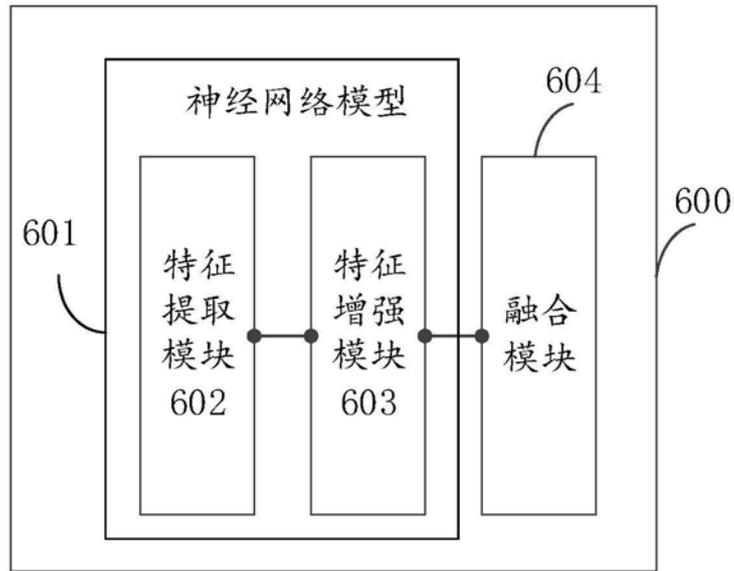


图10

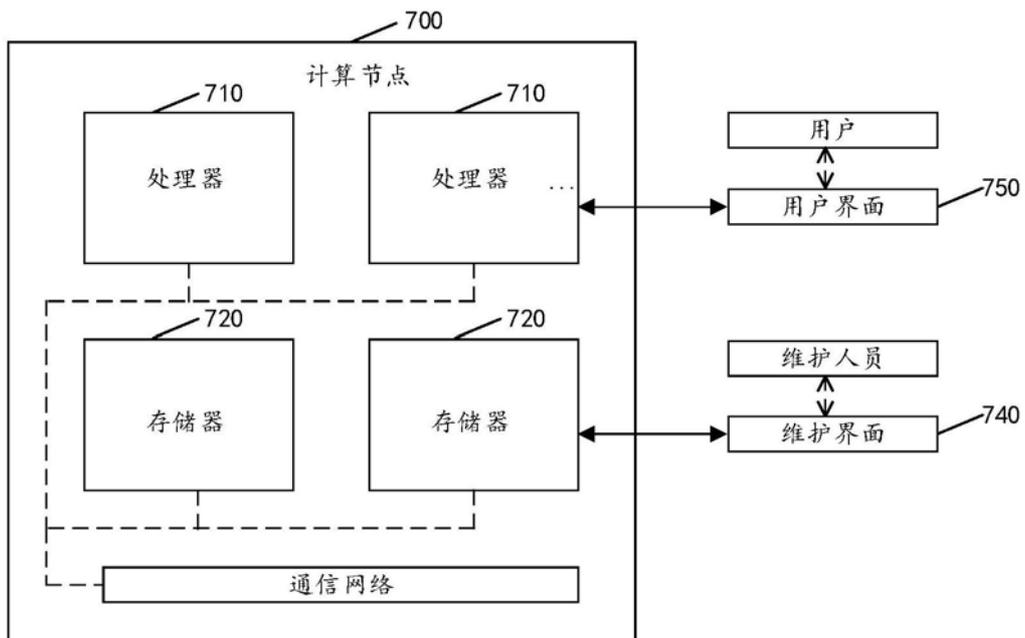


图11

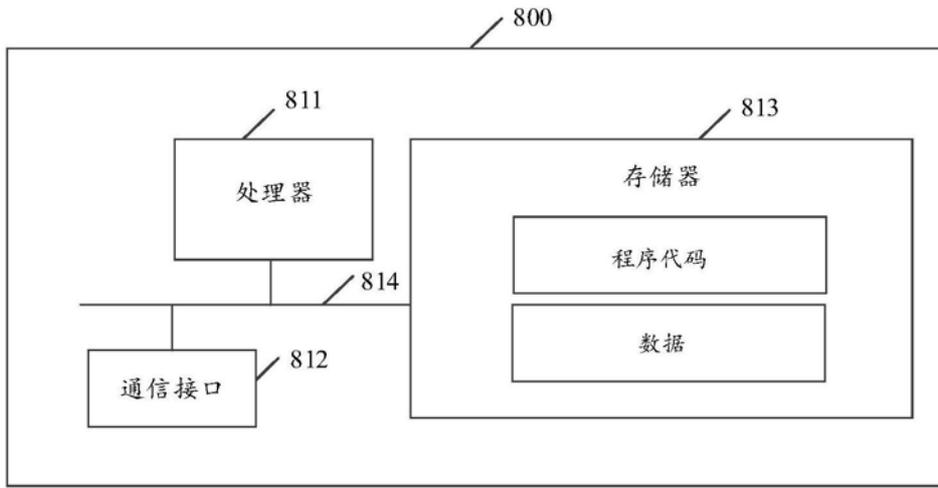


图12

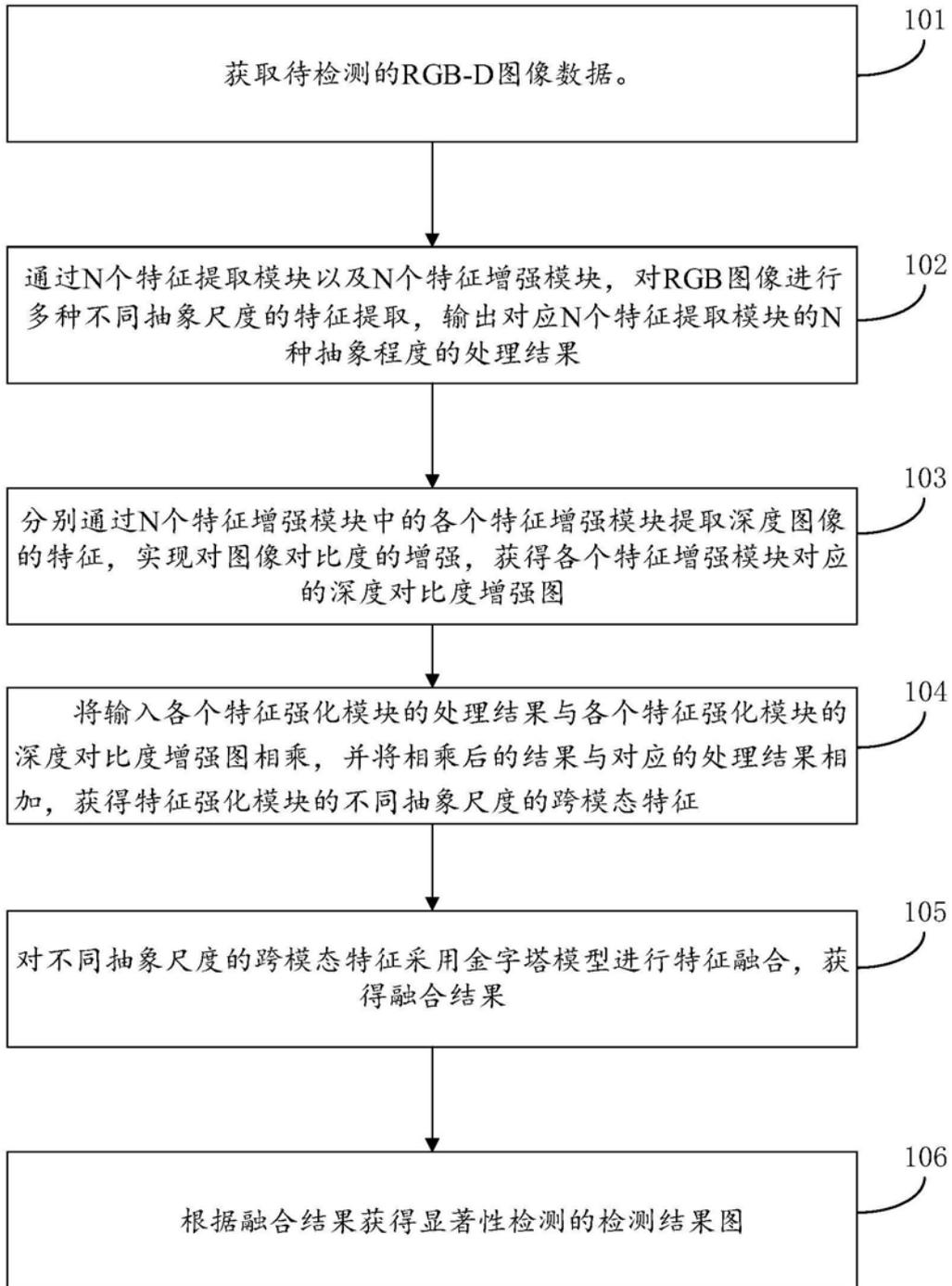


图13

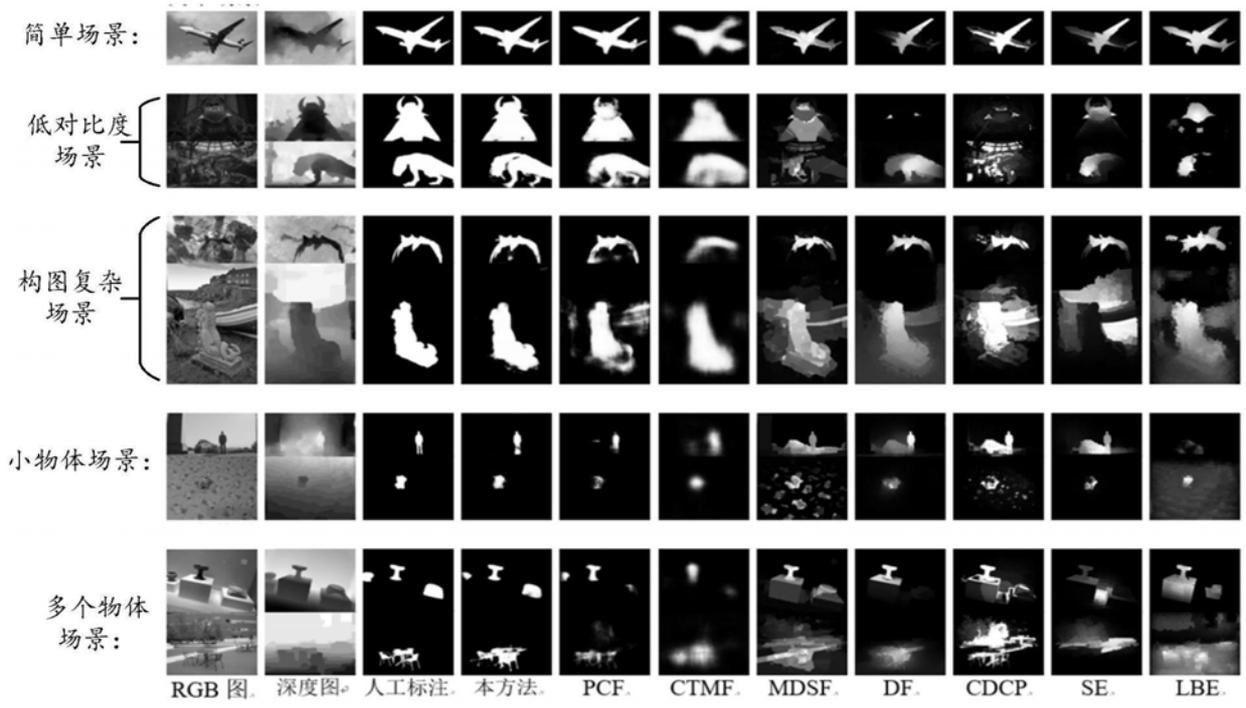


图14