



(12)发明专利申请

(10)申请公布号 CN 105809671 A

(43)申请公布日 2016.07.27

(21)申请号 201610119870.X

(22)申请日 2016.03.02

(71)申请人 无锡北邮感知技术产业研究院有限公司

地址 214135 江苏省无锡市新区菱湖大道97号大学科技园创新研发楼二期C楼

(72)发明人 马健翔 周瑜 宋桂岭

(74)专利代理机构 北京中恒高博知识产权代理有限公司 11249

代理人 宋敏

(51)Int.Cl.

G06T 7/00(2006.01)

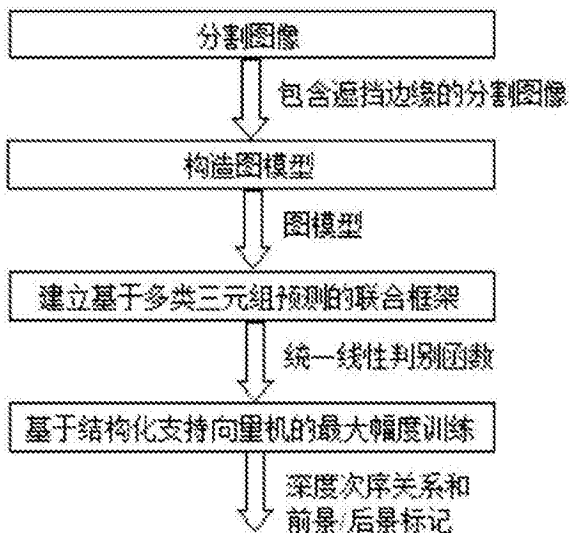
权利要求书2页 说明书7页 附图2页

(54)发明名称

前景区域标注与深度次序推理的联合学习方法

(57)摘要

本发明公开了一种前景区域标注与深度次序推理的联合学习方法,包括,分割图像,构造分割图像的图模型,建立基于多类三元组预测的联合框架,基于结构化支持向量机进行最大幅度训练,从而得出图像中各个区域的深度次序关系以及前后景标记。针对单目图像的深度估计和前景区域标记进行回顾和总结,并基于其单一推测的局限性和缺陷作出改进和优化,提出一种联合估计的框架,并且使用了Geometric Context数据集和Cornell Depth-Order数据集的部分图片验证算法的正确性和有效性。达到提升推测结果的有效性的目的。



1. 一种前景区域标注与深度次序推理的联合学习方法,其特征在于,包括以下步骤:

步骤1、分割图像,对图像进行物体级别的分割,即保留图像物体区域间的遮挡边界,并提取图像中区域和边界的特征向量;

步骤2、构造分割图像的图模型,即给定分割图像 X ,分割图像 X 是由 N 个区域 $\{x_i\}_{i=1}^N$ 和区域间边界 $E = \{(i, j) | \forall j \in N(i) \wedge i = 1, \dots, N\}$ 组成,将分割图像 X 抽象建立一个无向图模型,图的顶点代表区域,边表示需要被预测的深度次序,从而确定区域标签 $R = \{r_i | i = 1, \dots, N\} \in \{0, 1\}^N$ 和相对深度次序 $B = \{b_{ij} | \forall (i, j) \in E\} \in \{0, 1\}^{|E|}$;

步骤3、建立基于多类三元组预测的联合框架,为了衡量图模型表示的图像中区域间深度次序和几何关系的判别能力,基于步骤1提取的区域和边界的特征向量和步骤2构造的图模型,定义一个线性判别函数 $F(X, B, R; w, v)$;

$$\begin{aligned} F(X, B, R; w, v) &= \sum_{(i, j) \in E} D(x_i, x_j; w) b_{ij} + \sum_{i=1}^N E(x_i; v) r_i \\ &= \sum_{(i, j) \in E} \langle w, \phi_{ij}(x_i, x_j) \rangle b_{ij} + \sum_{i=1}^N \langle v, \phi_i(x_i) \rangle r_i \\ &= \langle \theta, \Phi(X, B, R) \rangle \end{aligned}$$

其中, w 和 v 为相应特征的权系数向量,然后训练一个把区域标签 R 和相对深度次序 B 联合起来的特征映射到联合标签的单一分类器,即定义一个三元组标签集 T ,并将线性判别函数进行更新为 $F(X, T; W)$;

步骤4、基于结构化支持向量机进行最大幅度训练,从而估计线性判别函数 $F(X, T; W)$ 中的权系数矩阵 W ,根据估计的权系数矩阵 W 得出图像中各个区域的深度次序关系以及前后景标记。

2. 根据权利要求1所述的前景区域标注与深度次序推理的联合学习方法,其特征在于,所述步骤1中的区域的特征向量中区域的特征,包括颜色、纹理、位置和形状,所述边界的特征向量中边界的特征,包括几何、凹凸性、位置和显著度。

3. 根据权利要求2所述的前景区域标注与深度次序推理的联合学习方法,其特征在于,步骤1中还包括,使用视觉词特征代表每一归属区域的后验概率。

4. 根据权利要求1所述的前景区域标注与深度次序推理的联合学习方法,其特征在于,步骤2中,对于区域标签 $R = \{r_i | i = 1, \dots, N\} \in \{0, 1\}^N$ 和相对深度次序

$B = \{b_{ij} | \forall (i, j) \in E\} \in \{0, 1\}^{|E|}$,其中 $b_{ij} = 0$ 代表区域 x_i 在 x_j 前面;否则, $b_{ij} = 1$;且当 x_i 为背景区域时, $r_i = 0$;否则, $r_i = 1$ 。

5. 根据权利要求4所述的前景区域标注与深度次序推理的联合学习方法,其特征在于,三元组标签集 $T = \{t_{ij} = (b_{ij}, r_i, r_j) | \forall i, j \in E\}$,

且三元组标签集的约束条件为:给定一个角点 $J = \{(i, j), (j, k), (k, i)\} \subset E$,由三个顶点 x_i, x_j 和 x_k 组成,则对所有角点 J 在边界标签上都要满足如下的环形不等式:

$$1 \leq \sum_{(i,j) \in J} b_{ij} \leq 2, \forall J \in \mathbf{J}$$

则线性判别函数为：

$$\begin{aligned} F(X, T; W) &= \sum_{(i,j) \in E} U(t_{ij}; x_i, x_j, W) \\ &= \sum_{(i,j) \in E} \langle W, \phi_{ij}^{joint}(x_i, x_j) \rangle \end{aligned}$$

其中, W 为权系数矩阵, $U(t_{ij}; x_i, x_j, W)$ 即为线性判别函数, $\phi_{ij}^{joint} = [\phi_{ij}, \psi_i, \psi_j]$, ϕ_{ij} 表示从分割图像提取出的边界特征向量, ψ_i, ψ_j 分别表示从区域 i 和 j 提取的区域特征向量。

前景区域标注与深度次序推理的联合学习方法

技术领域

[0001] 本发明涉及图像处理领域,具体地,涉及一种前景区域标注与深度次序推理的联合学习方法。

背景技术

[0002] 深度估计的研究一直是计算机视觉领域中基础又重要的问题。研究者们一开始着重于发掘图像的绝对深度次序,后来由于精确估计的困难性,并发现可以通过有效提取单目图像的深度线索,如遮挡和几何信息等,来推测物体之间的相对深度次序,并用于处理像显著性检测和三维场景理解等高级视觉问题。

[0003] 现有技术大多使用各种从轮廓和T角点结构提取出的局部线索来计算相对深度次序。然而,这一方法自然是有缺陷的,例如天空、地面这样的背景区域显然是在图像场景的最后面,而从边界和角点进行深度次序推断时却没有考虑图像场景区域的前景背景的布局特征,并且某一区域的几何信息有助于理解相邻区域间的相对深度布局。当估计区域标签时,成对区域的相对深度次序也提高了它的区域标签推测的正确性。D.Hoiem等(“Recovering occlusion boundaries from an image”,International Journal of Computer Vision,91(3):328-346,2011)的研究就得到了极大的关注。他们推测物体区域间的相对深度次序,并计算推测这一过程的场景布局置信度,对其分析效果。实验结果表明有了几何置信度线索的帮助,推测效果有了极大的改善。然而,一旦几何置信度被估计出来,相对深度次序的结果非常依赖于几何置信度的准确性。

[0004] 深度估计是计算机视觉领域中最有挑战性的问题之一。传统研究中大部分都聚焦在基于多视角的图像和运动线索估计精确深度值这一方面。然而,只依靠单目图像,这一任务很难完成。于是,研究者们使用各种从单目图像的轮廓和T角点结构提取出的局部线索来计算相对深度次序。然而,从角点和边界估计出的深度次序由于没有考虑到图像中区域前景背景的布局特征而具有天生的缺陷。

[0005] D.Hoiem等人的研究在传统单目线索的基础上添加了几何置信度线索,从而改善了推测效果。然而,一旦几何置信度被估计出来,相对深度次序的结果非常依赖于几何置信度的准确性。若几何置信度的偏差较大,则会严重影响最终的效果,使得系统缺乏较强的鲁棒性。深度次序估计和前景区域的标记是相辅相成的,然而现有的技术方法并未考虑到这一点,使得未能体现出它们的联系。

发明内容

[0006] 本发明的目的在于,针对上述问题,提出一种前景区域标注与深度次序推理的联合学习方法,以实现提升推测结果的有效性的优点。

[0007] 为实现上述目的,本发明采用的技术方案是:

[0008] 一种前景区域标注与深度次序推理的联合学习方法,包括以下步骤:

[0009] 步骤1、分割图像,对图像进行物体级别的分割,即保留图像物体区域间的遮挡边

界,并提取图像中区域和边界的特征向量;

[0010] 步骤2、构造分割图像的图模型,即给定分割图像X,分割图像X是由N个区域 $\{x_i\}_{i=1}^N$ 和区域间边界 $E = \{(i, j) \mid \forall j \in N(i) \wedge i = 1, \dots, N\}$ 组成,将分割图像X抽象建立一个无向图模型,图的顶点代表区域,边表示需要被预测的深度次序,从而确定区域标签 $R = \{r_i \mid i = 1, \dots, N\} \in \{0, 1\}^N$ 和相对深度次序 $B = \{b_{ij} \mid \forall (i, j) \in E\} \in \{0, 1\}^{|E|}$;

[0011] 步骤3、建立基于多类三元组预测的联合框架,为了衡量图模型表示的图像中区域间深度次序和几何关系的判别能力,基于步骤1提取的区域和边界的特征向量和步骤2构造的图模型,定义一个线性判别函数 $F(X, B, R; w, v)$;

[0012] 其中,w和v为相应特征的权系数向量,然后训练一个把区域标签R和相对深度次序B联合起来的特征映射到联合标签的单一分类器,即定义一个三元组标签集T,并将线性判别函数进行更新为 $F(X, T; W)$;

[0013] 步骤4、基于结构化支持向量机进行最大幅度训练,从而估计线性判别函数 $F(X, T; W)$ 中的权系数矩阵W,根据估计的权系数矩阵W得出图像中各个区域的深度次序关系以及前后景标记。

[0014] 优选的,所述步骤1中的区域的特征向量中区域的特征,包括颜色、纹理、位置和形状,所述边界的特征向量中边界的特征,包括几何、凹凸性、位置和显著度。

[0015] 优选的,步骤1中还包括,使用视觉词特征代表每一归属区域的后验概率。

[0016] 优选的,步骤2中,对于区域标签 $R = \{r_i \mid i = 1, \dots, N\} \in \{0, 1\}^N$ 和相对深度次序 $B = \{b_{ij} \mid \forall (i, j) \in E\} \in \{0, 1\}^{|E|}$,其中 $b_{ij} = 0$ 代表区域 x_i 在 x_j 前面;否则, $b_{ij} = 1$;且当 x_i 为背景区域时, $r_i = 0$;否则, $r_i = 1$ 。

[0017] 优选的,三元组标签集 $T = \{t_{ij} = (b_{ij}, r_i, r_j) \mid \forall i, j \in E\}$,

[0018] 且三元组标签集的约束条件为:给定一个角点 $J = \{(i, j), (j, k), (k, i)\} \subset E$,由三个顶点 x_i, x_j 和 x_k 组成,则对所有角点J在边界标签上都要满足如下的环形不等式:

$$1 \leq \sum_{(i,j) \in J} b_{ij} \leq 2, \forall J \in \mathbf{J}$$

[0019] 则线性判别函数为:

$$F(X, T; W) = \sum_{(i,j) \in E} U(t_{ij}; x_i, x_j, W)$$

$$= \sum_{(i,j) \in E} \langle W, \phi_{ij}^{joint}(x_i, x_j) \rangle$$

[0021] 其中,W为权系数矩阵, $U(t_{ij}; x_i, x_j, W)$ 即为线性判别函数, $\phi_{ij}^{joint} = [\phi_{ij}, \psi_i, \psi_j]$, ϕ_{ij} 表示从分割图像提取出的边界特征向量, ψ_i, ψ_j 分别表示从区域i和j提取的区域特征向量。

[0022] 本发明的技术方案具有以下有益效果:

[0023] 本发明技术方案针对单目图像的深度估计和前景区域标记进行回顾和总结,并基

于其单一推测的局限性和缺陷作出改进和优化,提出一种联合估计的框架,并且使用了Geometric Context数据集和Cornell Depth-Order数据集的部分图片验证算法的正确性和有效性。

[0024] 具体效果如下:

[0025] (1)提出了一种联合的深度次序推测和前景区域标记的共同学习框架,利用它们之间的关联信息,维持物理场景中近似真实的物体区域布局,从而提升推测结果的有效性;(2)建立一个无向图模型进行图像分割表达,并将为区域和边缘分配二值标签转化成带有约束的整数规划问题进行求解;(3)创新性地定义一个三元组变量用来描述相邻区域对以及之间的边缘,从而将问题转化成对多类别三元组量进行预测分类的问题,并且将原来不可解的硬约束规划问题转化成可解的软约束规划问题,从而简单有效的在全局上惩罚边界和区域标记推测不一致的情况。(4)由于定义的判别函数是线性的,因而使用结构化支持向量机技术进行参数学习,再用线性规划松弛化方法去求解,从而进行有效推测。

[0026] 下面通过附图和实施例,对本发明的技术方案做进一步的详细描述。

附图说明

[0027] 图1为本发明实施例所述的前景区域标注与深度次序推理的联合学习方法的流程图示意图;

[0028] 图2为本发明实施例所述的分割图像示意图;

[0029] 图3a和图3b为本发明实施例所述的无向图模型示意图。

具体实施方式

[0030] 以下结合附图对本发明的优选实施例进行说明,应当理解,此处所描述的优选实施例仅用于说明和解释本发明,并不用于限定本发明。

[0031] 如图1所示,本技术方案主要从以下四个步骤进行说明:分割图像,构造图模型,建立基于多类三元组预测的联合框架和基于结构化支持向量机的最大幅度训练。

[0032] 1.分割图像:

[0033] 首先,对图像进行物体级别的分割,也就是保留物体区域间的遮挡边界,并提取区域和边界特征向量用作之后的推测。一方面,区域特征包含颜色、纹理、位置和形状特征。除此以外,还使用视觉词特征来代表每一归属区域的后验概率。另一方面,边缘特征向量则用来估计深度次序,包括几何、凹凸性、位置和显著度特征。

[0034] 2.构造图模型:

[0035] 给定分割图像 X ,它是由 N 个区域和区域之间的边界组成。将其抽象建立一个无向图模型 $G=(V,E)$,图的顶点代表区域,边表示需要被预测的深度次序,从而确定区域标签 R 和相对深度次序 B 。

[0036] 3.建立基于多类三元组预测的联合框架:

[0037] 为了衡量图模型表示的图像中区域间深度次序和几何关系的判别能力,基于第一步骤提取的相应特征向量和第二步骤构造的图模型,定义一个线性判别函数 $F(X,B,R;w,v)$ 。

[0038] 在给定的图模型中,每一个二值标签必须满足物理上的近似真实条件,即三个相

邻区域两两之间的前景/背景关系是有效的,为此对区域间标签是有效深度次序关系进行数学化定义。另外,区域标签也必须和相对深度次序一致,从而前景区域保持在背景区域的前面。因此,需要联合估计,即训练一个能把它们联合起来的特征映射到联合标签的单一分类器,而不是对区域和边缘标签分别训练各自的分类器。为此,定义一个三元组标签集 T ,从而把问题转化成包含8个类别的多类分类问题,并将线性判别函数进行更新为 $F(X, T; W)$ 。

[0039] 既然深度次序必须物理上近似真实的,对三元组变量的约束应当遵循上述定义的有效深度次序关系。因而,寻求最佳三元组标签的最优化问题可进行公式化,并利用线性规划的松弛变量法进行求解,从而得出最优三元组标签 T^* ,因而最佳相对深度次序推测 B^* 可由最优三元组标签 T^* 直接确定出来。但是,区域标签 R 由于包含多种三元组变量组合情况而无法直接得出,于是通过多数投票机制来确定最优区域标签 R^* ,这一机制在经验上是有效的。

[0040] 4. 基于结构化支持向量机的最大幅度训练:

[0041] 这一步骤利用基于结构化支持向量机的技术进行大幅度的训练。为了估计线性判别函数 $F(X, T; W)$ 中的权系数矩阵 W ,参考I. Tsochantaridis等(“Large margin methods for structured and interdependent output variables”, JMLR, 6:1453-1484, 2005)利用真实标签 T_n 和预测标签 T 并基于汉明损失函数求解相应的约束最优化问题。而此最优化问题涉及到关于 T 维度的指数级约束。因此,使用割平面算法降低约束的数量级。从而,可定义出关于第 n 个训练数据的最违反约束标签,然后把它添加到约束集中。得出约束集后,最优化问题可由二次规划解决,从而求出最终结果。

[0042] 以下结合实际应用对本发明技术方案具体说明如下:

[0043] 分割图像:

[0044] 首先,对图像进行物体级别的分割,也就是保留物体区域间的遮挡边界,如图2所示,并提取区域和边界特征向量用作之后的深度推测。一方面,区域特征是一个52维的低层次的特征向量(D. Hoiem et al, “Recovering surface layout from an image”, IJCV, 75(1):151-172, 2007),包含颜色、纹理、位置和形状特征。除此以外,还使用150维的视觉词特征(D. Batra et al, “Learning class-

[0045] specific affinities for image labelling”, CVPR, 2008)来代表每一归属区域的后验概率。另一方面,边缘特征向量则用来估计深度次序,包括几何(4维)、凹凸性(2维)、位置(2维)和显著度特征(27维)。

[0046] 构造图模型:

[0047] 给定分割图像 X ,它是由 N 个区域 $\{x_i\}_{i=1}^N$ 和区域间边界

$E = \{(i, j) | \forall j \in N(i) \wedge i = 1, \dots, N\}$ 组成。将其抽象建立一个无向图模型 $G = (V, E)$,图的顶点代表区域,边表示需要被预测的深度次序,从而确定区域标签 $R = \{r_i | i = 1, \dots, N\} \in \{0, 1\}^N$ 和相对深度次序 $B = \{b_{ij} | \forall (i, j) \in E\} \in \{0, 1\}^{|E|}$ 。如图3a和图3b所示,其中 $b_{ij} = 0$ 代表区域 x_i 在 x_j 前面;否则, $b_{ij} = 1$ 。当 x_i 为背景区域时, $r_i = 0$;否则, $r_i = 1$ 。

[0048] 建立基于多类三元组预测的联合框架:

[0049] 为了衡量图模型表示的图像中区域间深度次序和几何关系的判别能力,定义如下线性判别函数:

$$\begin{aligned}
 F(X, B, R; w, v) &= \sum_{(i,j) \in E} D(x_i, x_j; w) b_{ij} + \sum_{i=1}^N E(x_i; v) r_i \\
 [0050] \quad &= \sum_{(i,j) \in E} \langle w, \phi_{ij}(x_i, x_j) \rangle b_{ij} + \sum_{i=1}^N \langle v, \varphi_i(x_i) \rangle r_i \\
 &= \langle \theta, \Phi(X, B, R) \rangle \tag{1}
 \end{aligned}$$

[0051] 其中, w 和 v 为相应特征的权系数向量, $D(x_i, x_j; w)$ 用来衡量哪一区域在前面, 得出较大的负值意味着区域 x_i 在前面, 反之较大正值意味着 x_j 在前面。类似地, $E(x_i; v)$ 若为负值则意味 x_i 为前景区域, 正值则为背景区域。 ϕ_{ij} 和 φ_i 分别表示从分割图像提取出的边缘特征向量和区域特征向量。

[0052] 在给定的图模型中, 每一个二值标签必须满足物理上的近似真实条件。三个相邻区域两两之间的适当前景/背景关系的可靠例子如表1所示, “0”意味着第 i 个分割区域在前面, “1”意味着第 j 个分割区域在前面。

分割		相对深度次序 b_{ij}						
x_i	x_j	inval	val	val	val	val	val	inval
x_1	x_2	0	0	0	0	1	1	1
x_2	x_3	0	0	1	1	0	0	1
x_3	x_1	0	1	0	1	0	1	0

[0054] 表1、环形深度次序的标签有效性表。

[0055] 区域间标签 b_{ij} 的约束可据如下数学化定义:

[0056] 定义1(有效的深度次序关系)给定一个角点 $J = \{(i, j), (j, k), (k, i)\} \subset E$, 它由三个顶点 x_i, x_j 和 x_k 组成, 则对所有角点 J 在边界标签上都要满足如下的环形不等式:

$$1 \leq \sum_{(i,j) \in J} b_{ij} \leq 2, \forall J \in \mathbf{J} \tag{2}$$

[0058] 除了上述约束, 区域标签 R 也必须和相对深度次序 B 一致, 从而前景区域保持在背景区域的前面。因此, R 和 B 必须联合估计, 即训练一个能把它们联合起来的特征映射到联合标签的单一分类器, 而不是对区域和边缘标签分别训练各自的分类器。为此, 定义一个三元组标签集 $T = \{t_{ij} = (b_{ij}, r_i, r_j) \mid \forall i, j \in E\}$ 。既然 r_i 和 b_{ij} 都是二值的, 那么 t_{ij} 则可以有8种不同的值, 从而对于它的推测可被考虑成包含8个类别的多类分类问题。加入这一约束后的判别函数可定义成:

$$\begin{aligned}
 F(X, T; W) &= \sum_{(i,j) \in E} U(t_{ij}; x_i, x_j, W) \\
 [0059] \quad &= \sum_{(i,j) \in E} \langle W, \phi_{ij}^{joint}(x_i, x_j) \rangle \tag{3}
 \end{aligned}$$

[0060] 其中, W 为权系数矩阵, $U(t_{ij}; x_i, x_j, W)$ 即为线性判别函数, 联合特征 ϕ^{joint} 就是把边缘特征和区域特征相连接形成, 即 $\phi_{ij}^{joint} = [\phi_{ij}, \psi_i, \psi_j]$ 。

[0061] 既然深度次序必须物理上近似真实的,对三元组变量的约束应当遵循等式(2)。因而,寻求最佳三元组标签的最优化问题可公式化为:

$$\begin{aligned}
 & T^* = \arg \max_{T \in \mathbf{T}} F(X, T; W), \\
 [0062] \quad & s.t. \quad 1 \leq \sum_{(m,n) \in J} b_{mn}(t_{mn}) \leq 2, \quad \forall J \in \mathbf{J}
 \end{aligned} \tag{4}$$

[0063] 这里,

$$[0064] \quad b_{mn}(t_{mn}) = \begin{cases} 0 & \text{if } t_{mn} \geq 4 \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

[0065] 利用线性规划的松弛变量法,等式(4)可通过如下方式求解:

$$\begin{aligned}
 & Z^* = \arg \max_Z \sum_{(i,j) \in E} \sum_{t_{ij}=0}^7 \langle W, \phi_{ij}^{joint}(x_i, x_j) \rangle z_{ij}^{t_{ij}}, \\
 [0066] \quad & s.t. \quad \sum_{t_{ij}=0}^7 z_{ij}^{t_{ij}} = 1, \quad 0 \leq z_{ij}^{t_{ij}}, \\
 & 1 \leq \sum_{(m,n) \in J} \left[\sum_{t_{mn} \geq 4} z_{mn}^{t_{mn}} + \sum_{t_{mn} < 4} (1 - z_{mn}^{t_{mn}}) \right] \leq 2, \\
 & \forall J \in \mathbf{J}, t_{ij} = 0, \dots, 7
 \end{aligned} \tag{6}$$

[0067] 这里,

$$[0068] \quad Z = \begin{bmatrix} z_{12}^0 & \dots & z_{12}^7 \\ \vdots & \ddots & \vdots \\ z_{N-1,N}^0 & \dots & z_{N-1,N}^7 \end{bmatrix} \tag{7}$$

[0069] 得出的Z*即为最优三元组标签T*,并且t_{ij}和b_{ij}之间是一一对应的,因而B*可由T*直接确定出来。但是,R由于包含多种三元组变量组合情况而无法直接得出,于是通过多数投票机制来确定区域标签R*,这一机制在经验上是有效的。

[0070] 4. 基于结构化支持向量机的最大幅度训练

[0071] 这一步骤利用基于结构化支持向量机的技术进行大幅度的训练。为了估计线性判别函数F(X,T;W)中的权系数矩阵W,参考I.Tsochantaridis等的研究方法(“Large margin methods for structured and interdependent output variables”,JMLR,6:1453-1484,2005)求解如下的约束最优化问题:

$$\begin{aligned}
 & \min_{w, \xi} \frac{1}{2} \|w\|^2 + \frac{C}{N} \sum_{n=1}^N \xi_n, \\
 [0072] \quad & s.t. \quad d(X_n, T; W) \geq \Delta(T_n, T) - \xi_n, \quad T \in \mathbf{T} \setminus T_n, \forall n, \\
 & \xi_n \geq 0, \quad \forall n
 \end{aligned} \tag{8}$$

[0073] 其中,d(X_n,T;W)表示真实标签T_n和预测标签T之间的判别函数值的差,例如

[0074] $d(X_n, T; W) = F(X_n, T_n; W) - F(X_n, T; W)$ (9)

[0075] ξ_n 是用来允许 X_n 存在训练误差的松弛变量, C 是用来权衡训练误差最小化和幅度最大化的平衡系数。损失函数 $\Delta(T_n, T)$ 用来在给定正确标签 T_n 的情况下对预测标签 T 进行误差衡量。这里利用的是改进的汉明损失函数(S.Kim et al, “Task-specific image partitioning”, 2012)。

[0076] 公式(8)的最优化问题涉及到关于 T 维度的指数级约束。因此, 使用割平面算法降低约束的数量级。从而, 关于第 n 个训练数据的最违反约束标签可由如下定义式得出:

$$[0077] \quad \hat{T}_n = \arg \max_{T \in \mathcal{T} \setminus T_n} [\Delta(T_n, T) - d(X_n, T; W)] \quad (10)$$

[0078] 然后把它添加到约束集中。这里考虑的损失函数可在测试边缘上分解从而对 \hat{T}_n 进行有效推测。得出约束集后, 最优化问题可由二次规划解决, 从而求出最终结果。

[0079] 综上所述, 本发明的有益效果具体可通过表2看出:

[0080] 表2列出了是本发明方法和其余两种方法进行前景/后景推理的比较结果, 结果的数值越大表示推理的结果越准确。另外对于推理前的准备步骤, 即图像分割, 本发明技术方案采用了两种不同的分割算法: “Seg-ho”为D.Hoiem等使用的图像分割算法, 而“GT seg”则表示Z.Jia等(“A learning-based framework for depth ordering”, CVPR, 2012, pp:294-301)使用的基于真实标记的分割。从表2中可以看出, 本发明方法都取得了优于其它方法的最高准确率。其中(I)D.Hoiem等提出的方法, (II)Z.Jia等提出的方法, (III)本技术方案提出的方法。

方法	Geometric Context 数据集		D-order 数据集
	Seg-ho	GT seg	GT seg
(I)	79.9	-	-
(II)	-	73.3	91.7
(III)	81.5	80.8	93.2

[0081] 表1基于Geometric Context和Cornell Depth-Order数据集的前景/后景推理的准确率。

[0083] 最后应说明的是: 以上所述仅为本发明的优选实施例而已, 并不用于限制本发明, 尽管参照前述实施例对本发明进行了详细的说明, 对于本领域的技术人员来说, 其依然可以对前述各实施例所记载的技术方案进行修改, 或者对其中部分技术特征进行等同替换。凡在本发明的精神和原则之内, 所作的任何修改、等同替换、改进等, 均应包含在本发明的保护范围之内。

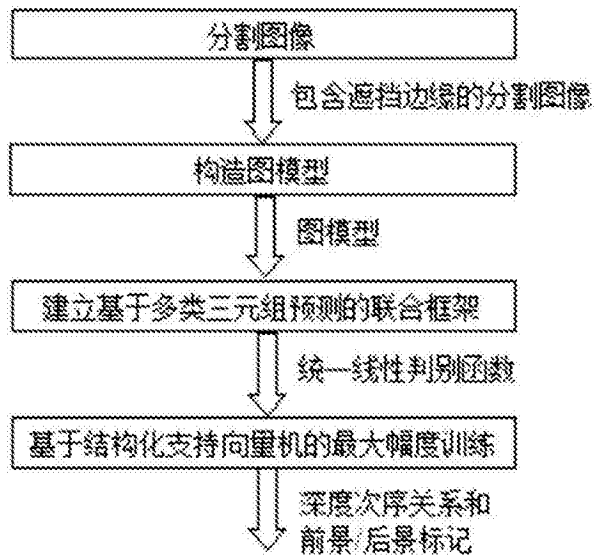


图1

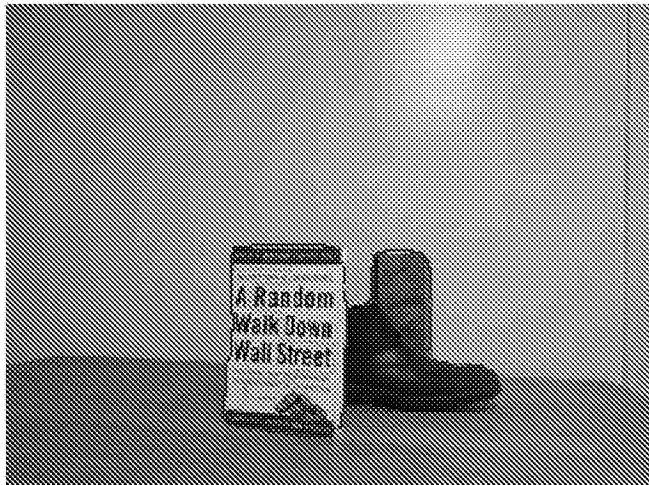


图2

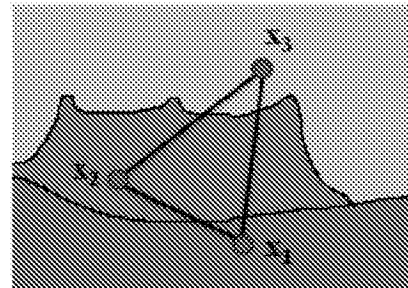


图3a

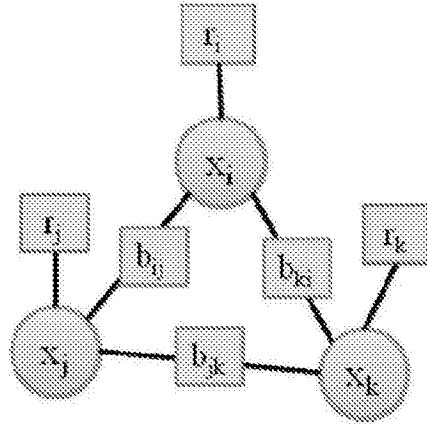


图3b