

一种组合型损失函数的位姿估计算法

张德, 李国璋, 王怀光*, 张峻宁

陆军工程大学石家庄校区车辆与电气工程系, 河北 石家庄 050003

摘要 基于卷积神经网络(CNN)的目标位姿估计模型的损失函数大多采用两点之间的欧氏距离作为评判准则,虽然该损失函数计算简单、运算速度快,但训练规则不够全面、缺乏对目标的全局认识。针对这一问题,提出了一种基于组合型损失函数的 ComPoseNet 模型,并进行位姿估计。此模型中的损失函数从空间学习的角度出发,同时利用两点欧氏距离、两点直线和两点直线角度等作为训练规则。相比传统损失函数,此算法分别从点、线以及角度方面考虑了目标的空间整体位置,进一步减小了估计位姿与真实位姿之间的误差,位姿估计得以改善。在 LineMod 数据上进行大量的实验和分析,结果表明,在相同的训练次数情况下,本文算法比传统算法收敛速度快、精度高、误差小,其中平移误差降低了 7.407%,角度误差降低了 6.968%。

关键词 图像处理; 损失函数; 位姿估计; 深度学习; 神经网络; 计算机视觉

中图分类号 TP399

文献标识码 A

doi: 10.3788/LOP56.221002

Pose Estimation Algorithm Based on Combined Loss Function

Zhang De, Li Guozhang, Wang Huaiguang*, Zhang Junning

Department of Vehicle and Electrical Engineering, Army Engineering University Shijiazhuang Campus,
Shijiazhuang, Hebei 050003, China

Abstract The loss function of a target pose estimation model based on a convolutional neural network (CNN) mostly uses the Euclidean distance between two points as the evaluation criterion. Although the loss function is simple in calculation and fast in operation, the training rules are not comprehensive enough and lack global understanding of the target. In this paper, a ComPoseNet model based on a combined loss function is proposed for pose estimation. The loss function in this model is based on spatial learning, and the two-point Euclidean distance, straight line, and straight line angle are used as training rules. Compared with the traditional loss function, this algorithm considers the spatial position of the target from the point, line, and angle, reducing the error between the estimated and the real poses so that the effect of the pose estimation is improved. Numerous experiments and analysis of LineMod data show that the algorithm has a higher convergence speed, greater accuracy, and smaller errors than the traditional algorithm operating for the same training times. The translation error is reduced by 7.407%, and the angle error is reduced by 6.968%.

Key words image processing; loss function; pose estimation; deep learning; neural network; computer vision

OCIS codes 100.3008; 100.4996; 150.1135

1 引言

近年来,随着计算机设备的更新换代以及神经网络^[1-4]理论的发展,深度学习即深层神经网络模拟人类大脑,利用视觉接受信息,目前在位姿估计^[5]方面备受青睐。

但目前位姿估计算法普遍存在计算精度低、位姿估计效果差等问题,其主要原因是参数调整不够精确、模型收敛精度不够高、模型的位姿输出结果与真实位姿相差较大。其中参数的调整是根据损失函数^[6]最小值进行的,模型收敛时损失函数值越小,参数调整越精确,位姿估计效果越好。由上述分析可

收稿日期: 2019-03-04; 修回日期: 2019-03-30; 录用日期: 2019-05-15

基金项目: 国家自然科学基金(51205405, 51305454)

* E-mail: 654959514@qq.com

知,损失函数的类型和最小值直接影响模型参数的准确性,进而影响模型的收敛精度和位姿估计效果。

传统的损失函数^[7-8]大多采用两点之间的欧氏距离作为计算规则。在二维坐标系中利用两点在 x 、 y 方向上的差值构造损失函数,在损失函数最小化过程中完成参数调整,最终模型达到收敛,进而完成目标检测和识别。此类传统损失函数在简单的二维平面中目标识别效果好、准确率高;但当应用在三维目标的位姿测量时,由于传统损失函数仅考虑点与点之间的差异,未从空间整体上对目标进行综合考虑,其位姿估计效果较差。这是因为二维平面中的目标识别仅涉及平面坐标 x 、 y 两个值,采用点与点之间的欧氏距离作为损失函数可以准确检测目标物、识别目标物;三维空间中的目标物位姿涉及立体空间,仅采用点与点之间的相关测量不能准确估计目标位姿,需增加更加全面的约束条件,并进行严格的计算与测量。为此,许多学者提出了多种网络模型结构和损失函数。如 Do 等^[9]在传统损失函数的基础上采用解耦方式输出 6D 姿态,利用李代数 so(3)转换为 3D 旋转矩阵后进行训练,同时利用 2D 边界框中心得到平移向量中的 x 、 y ,故此模型只需预测 z 分量和旋转向量 3 个值即可,减少了位姿估计过程中的计算量。Wang 等^[10]在传统损失函数的基础上,采用 RGB 图避开了用 2D 信息估计 3D 姿态时精度低的问题,同时利用 3D CAD 模型对深度图的不敏感性提升模型性能,这是一种基于 RGB 图片和 3D CAD 模型提升姿态精度的尝试。Li 等^[11]提出了一种新的基于目标 6D 姿态估计的深层神经网络结构。在给出初始姿态的基础上,通过对合成 RGB 图像渲染实现目标输入图像的匹配,在此基础上优化出更准的目标姿态。Jafari 等^[12]利用几何优化预测 6D 位姿,提出第一个基于深度学习且可以估计部分遮挡位姿的深层神经网络模型,使用实例分割网络消除了混乱背景和遮光板的影响,通过将像素密集地映射到 3D 对象表面位置,消除了对象外观的影响。Krull 等^[13]提出了能大幅度降低梯度方差且不显著增加计算时间的损失函数。通过离散选择提高效率,采用强化学习的随机估计系统,以端到端的方式训练,优化了现有的最先进的姿态参数。

上述几种模型和损失函数虽然在一定程度上提高了位姿估计的速度和准确度,但仅对计算能力和模型结构进行了改进,未对计算规则进行创新。本文从计算准则方面对传统损失函数进行了改进,从全局和局部角度出发,分别将两点欧氏距离、两点直

线、直线角度和多线约束损失函数作为约束条件进行位姿估计;利用此组合型损失函数^[14-16]进行训练,训练规则更符合位姿测量的要求,提高了模型的收敛速度和精度,降低了位姿估计的误差,后续实验也证实了该组合型损失函数的优越性。

2 基本原理

2.1 神经网络模型

本文算法 ComPoseNet 模型采用卷积神经网络^[17](CNN)作为基础框架,其主要特点为:1)该模型中的神经元之间非全连接;2)神经元之间的权重参数共享。这降低了网络结构的复杂度,减少了权重参数的个数,提高了模型的计算速度,节省了模型的运算时间。

ComPoseNet 模型以 CNN 网络结构为基础,使用 RGB 图像作为输入数据,经过神经网络运算得到目标物角点的二维投影,即 x 、 y 坐标,同时利用 PnP 原理回归得到目标物的三维信息,即目标位姿,这是一种端到端的算法,具体模型结构见图 1。图 1(a)表示位姿估计的整个运算过程和整体框架;图 1(b)单独展示了预处理模型和 ComPoseNet 模型的参数和具体结构,其中预处理采用大型数据集 PASCAL VOC 进行预训练,防止学习过程中出现网络对目标所在背景的过拟合现象;图 1(c)定性介绍了图像在运算过程中的形象变化,同时定量阐述了不同阶段图像的尺寸参数。

ComPoseNet 模型共 25 个卷积层、5 个池化层,卷积核和池化层具体参数见图 1(c),通过卷积池化运算将图像分割为 $416 \times 416 \times 32$, $208 \times 208 \times 32$, \dots , $13 \times 13 \times 1024$ 张量,最终分割成 $13 \times 13 \times 18$ 的三维张量。其中 13×13 表示将输入图像划分为 13×13 二维网格,目的是减少目标检测过程中的对比范围,仅在目标物中心点所在的网格进行接下来的位姿回归优化运算;18 代表每个探测目标相关联的多维向量,此向量包括 8 个角点的 x 、 y 坐标、目标的类别和置信度,即 $8 \times 2 + 1 + 1 = 18$ 。利用上述得到的二维数据信息结合 PnP 法则,最终构建出目标位姿,即三维矩阵 \mathbf{R} 和三维向量 \mathbf{t} 。在测试阶段,对比利用组合型损失函数得到的位姿与目标的真实位姿,从而优化、调整模型参数,最终得到误差最小的目标最优估计位姿。

2.2 数据集

采用 LineMod 数据集集中的 ape 类,此数据集是 6D 位姿估计中的标准数据集。包括大约 1200 张图像,其中训练集图像约 200 张,测试集图像约 1000 张。

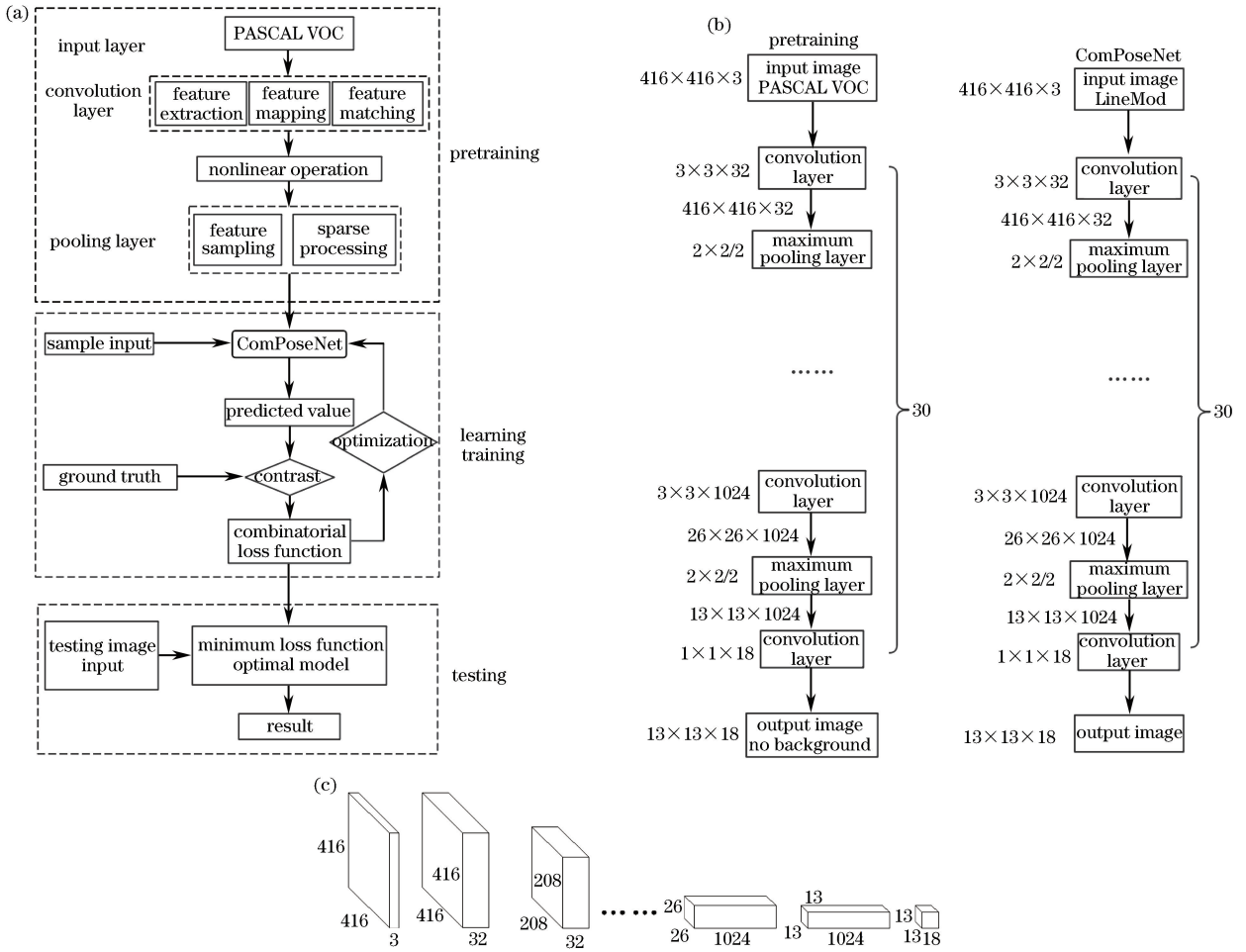


图 1 ComPoseNet 模型。(a) ComPoseNet 工作流程；(b)预处理和 ComPoseNet 结构；(c)图像处理过程
Fig. 1 ComPoseNet model. (a) ComPoseNet workflow; (b) preprocessing and ComPoseNet architecture; (c) image processing

3 改进的损失函数

损失函数表示预测值与真实值的差异程度,作用是根据估计值衡量模型的好坏。损失函数的一般表达式为

$$x_{\text{loss}} = F(y_j, y), \quad (1)$$

式中: y_j 为模型输出结果即预测值; j 为模型输出个数; y 为理想输出结果; $F(\cdot)$ 为计算函数; x_{loss} 为损失函数值。损失函数的具体意义和作用见图 2。

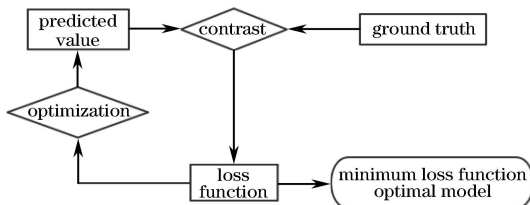


图 2 损失函数的作用
Fig. 2 Effect of loss function

由图 2 可知,在预测值和真实值对比过程中,通过调整参数优化模型,损失函数最终达到最小值,即预测值与真实值相差最小,模型达到最优。

3.1 传统损失函数

传统的损失函数采用两点之间的欧氏距离作为衡量标准,利用反向传播过程将欧氏距离最小化,得到最优估计模型。典型的传统损失函数主要有两种:平均绝对误差(MAE)和均方误差(MSE)。

3.2 组合型损失函数

传统损失函数将目标拆分为单个点,未从整体考虑,算法不够全面,效果未达到最佳。鉴于此,本研究提出一种考虑更全面的损失函数^[18-22],利用点与点之间构成的直线夹角 θ 以及两直线间夹角 α ,结合上述欧氏距离 d 作为评判标准,如图 3 所示。

图 3(a)为几何角点和直线的选取规则,选取 8 个角点作为基本点,分别为 O_1, O_2, \dots, O_8 ;定义角

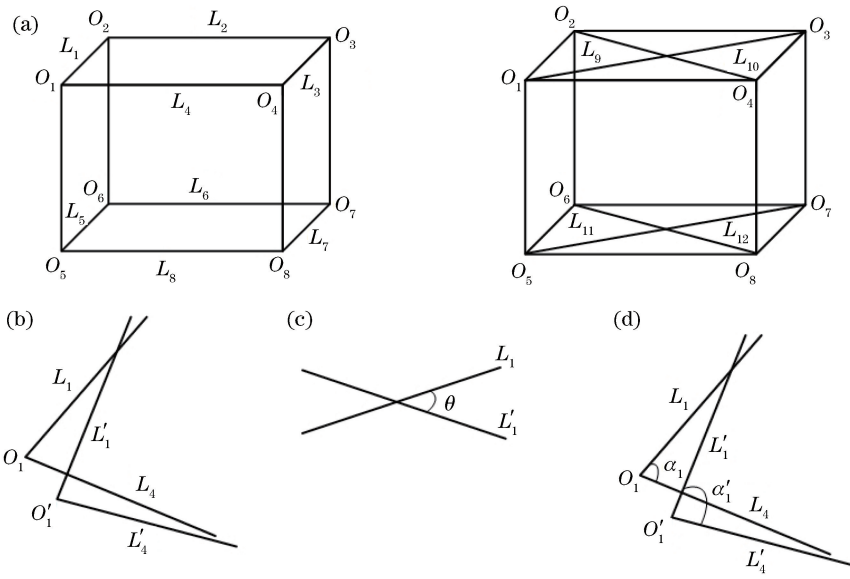


图3 几何规则示意图。(a)选取规则示意图;(b)距离 d ; (c)夹角 θ ; (d)位置角 α

Fig. 3 Geometric rule diagrams. (a) Selection rule schematic; (b) distance d ; (c) angle θ ; (d) angle α

点 O_1 和角点 O_2 之间的连线为直线 L_1 , 依次类推, 形成 8 条直线, 分别为 L_1, L_2, \dots, L_8 ; 此外, 定义角点 O_1 和角点 O_3 之间对角线为直线 L_9 , 依次类推, 形成 4 条直线, 分别为 $L_9, L_{10}, \dots, L_{12}$ 。图 3(b) 为目标真实角点 O_1 与估计角点 O'_1 在 x, y 方向上的欧氏距离 d_x, d_y ; 图 3(c) 为目标两真实角点之间直线 L_1 与两估计角点之间直线 L'_1 间的夹角 θ ; 图 3(d) 为两真实直线 L_1, L_2 的夹角 α_1 与两估计直线 L'_1, L'_2 的夹角 α'_1 。

本文算法计算规则表达式为

$$L(d, \theta, \alpha) = \begin{cases} (1 - \omega)L(d_x) + \omega L(\theta) + \omega L(\alpha) \\ (1 - \omega)L(d_y) + \omega L(\theta) + \omega L(\alpha) \end{cases}, \quad (2)$$

式中: ω 为超参数, 用来调整损失函数中距离和角度之间的比重; $L(d_x), L(d_y), L(\theta), L(\alpha)$ 分别表示 x, y, θ, α 方向的损失函数。

结合图 3 可知, 本文角点共有 8 个, 相应的 d_x, d_y 及 α 的个数为 8 即 $n = 8$; 直线有 12 条, 对应的 θ

有 12 个即 $m = 12$; i, j 表示各参量的数量; 具体表达式为

$$L(d_x) = \frac{1}{n} \sum_{i=1}^n (x_i - x)^2, \quad (3)$$

在 x 方向上, x_i 为模型输出结果, x 为理想输出结果;

$$L(d_y) = \frac{1}{n} \sum_{i=1}^n (y_i - y)^2, \quad (4)$$

在 y 方向上, y_i 为模型输出结果, y 为理想输出结果;

$$L(\theta) = \frac{1}{m} \sum_{i=1}^m (\theta_i - \theta)^2, \quad (5)$$

在 θ 角度上, θ_i 为模型输出结果, θ 为理想输出结果;

$$L(\alpha) = \frac{1}{n} \sum_{i=1}^n (\alpha_i - \alpha)^2, \quad (6)$$

在 α 角度上, α_i 为模型输出结果, α 为理想输出结果。

采用本文组合型损失函数和传统损失函数分别

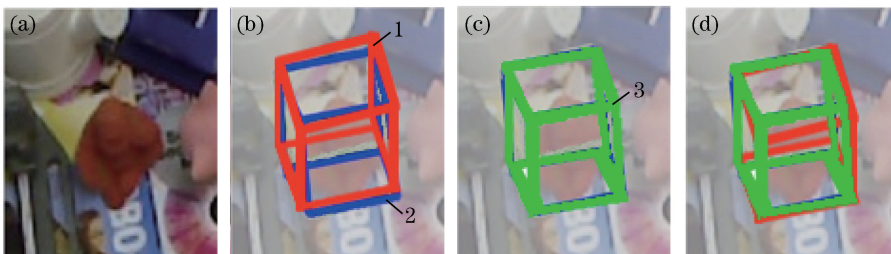


图4 不同损失函数的效果。(a)原图;(b)传统法;(c)本文算法;(d)对比

Fig. 4 Effects of different loss functions. (a) Original image; (b) traditional method; (c) proposed algorithm; (d) comparison

图 4 中 2 号框为目标的真实位姿,1 号框为使用传统损失函数估计的位姿,3 号框为本文算法估计得到的位姿。可以看出,1 号框比 2 号框大很多,贴合不够紧密;3 号框和 2 号框大小基本一致,紧密重合在一起。以上分析可以定性地看出,本文算法 ComPoseNet 的位姿估计效果明显优于传统损失函数算法。

4 实验

4.1 实验过程

本研究引入大型数据集 PASCAL VOC,对 ComPoseNet 模型进行预训练,减少了 RGB 图片背景在学习过程中对模型收敛的影响。数据集是 LineMod 中的 ape 类,属性为 RGB 图像,利用大约 1000 张图像对模型进行训练,调整模型组合型损失函数中的参数;利用 200 张左右图片对新模型的收敛效果进行测试。训练轮数超参设置为 200,每隔 5 epoch 进行一次测试。评估标准采用位姿估计中常用的 4 个变量,分别为像素、平移、角度误差和准确度。此外,实验引入传统法,与本文算法在收敛速度和收敛精度上进行对比,验证 ComPoseNet 的效果和性能;同时引入 SSD(Single Shot MultiBox Detector)和 YOLO(You Only Look Once)两种当前性能比较好的位姿估计算法作对比,衡量本文算法的效果。

实验配置为 CPU i5-8300H、显卡 GTX 1050Ti、内存 8 G。

4.2 实验结果

图 5、6 为采用不同算法得到的结果。由图 5、6 可见,本文算法曲线比传统算法起点高,在训练过程中,本文算法比传统算法收敛速度快、平衡点低、收

表 1 不同损失函数下的误差大小和准确率

Table 1 Errors and Accuracies of different loss functions

Loss function	Number of training rounds	Translation error /m	Angle error /($^{\circ}$)	Pixel error	Accuracy /%
Traditional	200	0.0459	8.310	5.982	86.391
Proposed	200	0.0425	7.731	5.576	91.262

此外,为验证本文的位姿估计效果,将本文算法和其他改进算法作对比,如图 7 所示。结果表明,本文算法在位姿估计效果上可与当前流行的 SSD^[23]和 YOLO^[24]算法相媲美,甚至比当前流行的算法效果好,详细数据见表 2。

4.3 损失函数的研究

4.3.1 各参数的作用

为详细研究本文算法中各参数的意义,另设计实验验证直线夹角 θ 、两直线间夹角 α 在组合型损

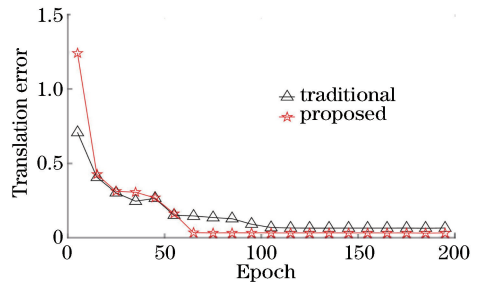


图 5 不同算法的平移误差

Fig. 5 Translation errors of different algorithms

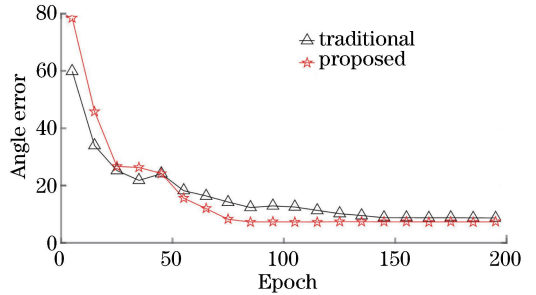


图 6 不同算法的角度误差

Fig. 6 Angle errors of different algorithms

敛精度高。可见,模型在学习初始阶段距离误差小,而角度误差大,并且仅依据距离误差作为学习准则,模型收敛精度低,达不到理想的位姿估计效果。本文算法综合点距离、单线夹角、多线角约束作为训练规则,在收敛精度方面比传统法有较大提高,具体结果如表 1 所示。

从表 1 可以看出,在训练轮数为 200 的情况下,本文算法的像素、平移和角度误差均比传统损失函数算法小,准确率比传统算法高,其中像素误差降低了约 6.787%,平移误差降低了约 7.407%,角度误差下降了约 6.968%,准确率提高了 5.638%。

失函数中的作用,结果如图 8、9 所示。其中三角形曲线表示距离 d 、直线夹角 θ 作为约束;五角星曲线表示距离 d 、直线夹角 α 作为约束;六角星曲线表示距离 d 、直线夹角 θ 、两直线夹角 α 同时作为约束。

由图 8、9 可见,三角形曲线平衡位置最低、五角星曲线居中,六角星曲线最高。可见,直线夹角 θ 、两直线间夹角 α 作为约束准则提高了模型的收敛精度,降低了模型的平移和角度误差,增强了位姿估计效果。估计效果如图 10 所示。

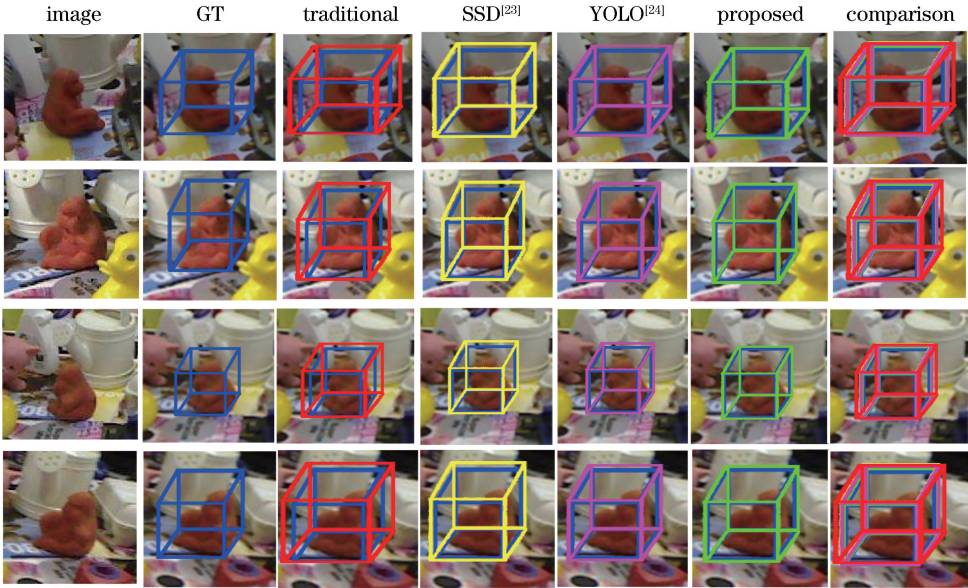


图 7 不同算法的位姿估计效果

Fig. 7 Effects of pose estimation of different algorithms

表 2 不同算法的误差

Table 2 Errors of different algorithms

Algorithm	Number of training rounds	Translation error / m	Angle error / (°)
Traditional	200	0.0459	8.310
SSD	200	0.0437	7.925
YOLO	200	0.0431	7.889
Proposed	200	0.0425	7.731

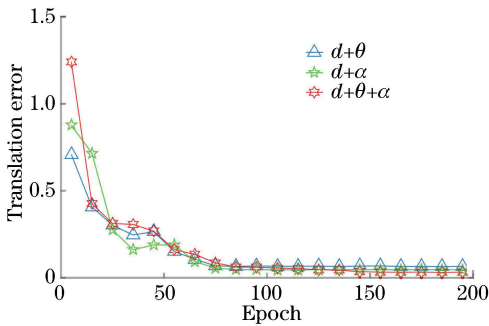


图 8 不同参数对平移误差的影响

Fig. 8 Effects of different parameters on translation error

由图 10 可见,1 号框最大,其次为 4 号框,再次为 5 号框,3 号框最小,夹角 θ 和夹角 α 作为约束准则提高了模型收敛精度,定量结果如表 3 所示。

从表 3 可以看出,仅采用距离 d 作为约束规则时,位移误差和角度误差最大,而随着约束条件的增加,模型收敛效果变好,误差减小;其中夹角 θ 、夹角 α 分别和距离 d 结合作为训练规则时,模型收敛精度也有一定提高,距离 d 同时结合夹角 θ 和夹角 α

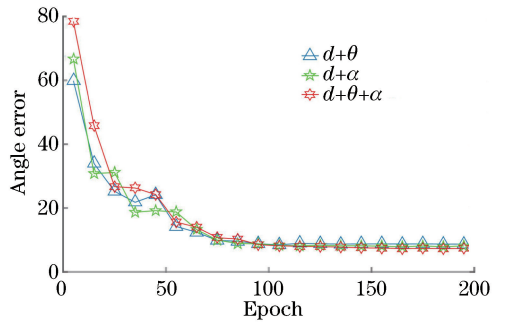


图 9 不同参数对角度误差的影响

Fig. 9 Effects of different parameters on angle error

组成组合型损失函数时,效果最佳。

4.3.2 损失函数的普适性

为验证此损失函数在其他方面的适用性,另设计实验将此组合型损失函数应用到目标检测领域,具体结果如图 11 所示。

图 11 中,2 号框为目标真实的检测效果,3 号框为采用本文算法得到的效果。可定性得出,本文算法在目标检测方面也具有一定的适应性,效果较好。

5 分析与讨论

传统损失函数算法未考虑目标的几何整体性,仅采用两点之间的欧氏距离进行运算,此过程虽然计算简单、初始值小,但收敛速度慢、最终收敛值较大、学习效率低、模型效果不佳。本文算法提出的组合型损失函数在考虑了传统损失函数的缺陷后,

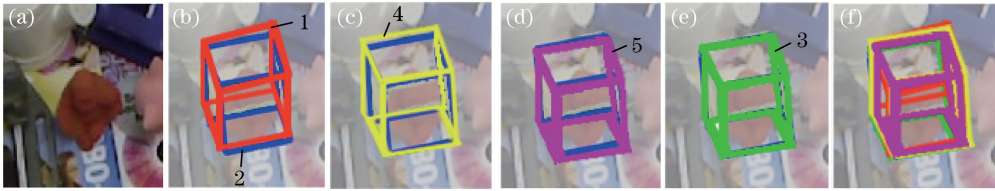


图 10 不同参数约束的估计效果图。(a)目标物;(b) d ;(c) $d+\theta$;(d) $d+\alpha$;(e) $d+\theta+\alpha$;(f)对比
Fig. 10 Estimation effect diagrams of different parameter constraints. (a) Object; (b) d ;
(c) $d+\theta$; (d) $d+\alpha$; (e) $d+\theta+\alpha$; (f) comparison

表 3 不同参数下的误差大小

Table 3 Errors under different parameters

Loss function	Number of training rounds	Translation error /m	Angle error /($^{\circ}$)
Traditional	200	0.0459	8.310
Proposed ($d+\theta$)	200	0.0442	8.125
Proposed ($d+\alpha$)	200	0.0439	7.984
Proposed ($d+\theta+\alpha$)	200	0.0425	7.731

利用两点直线、直线角度,多线约束的损失函数作为学习准则,同时兼顾两点之间的欧氏距离,构造的 ComPoseNet 模型收敛速度快、收敛值小、学习效率

高、模型效果好。

6 结 论

结合当前损失函数存在的不足,提出了一种新的组合型损失函数算法 ComPoseNet。在两点约束的基础上,采用两点直线、直线角度,多线作为组合型损失函数中的另外部分,考虑了目标整体结构,提高了模型的收敛精度,达到了较好的效果。在 LineMod 数据集上的实验结果表明,本文算法比传统算法收敛速度快、收敛精度高,其中平移误差降低了约 7.407%,角度误差下降了约 6.968%。

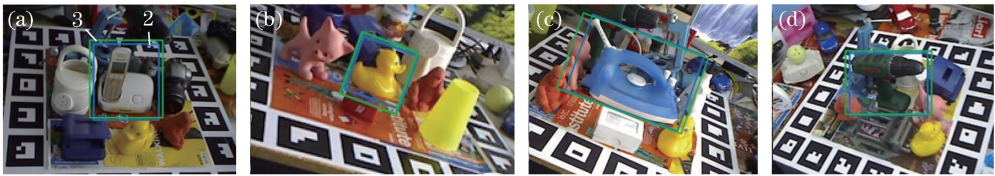


图 11 目标检测效果。(a)手机;(b)鸭子;(c)熨斗;(d)电钻
Fig. 11 Target detection effects. (a) Telephone; (b) duck; (c) iron; (d) drill

参 考 文 献

[1] Wu A C. Neural network and deep learning [M]. Beijing: Publishing House of Electronics Industry, 2016.
吴岸城. 神经网络与深度学习[M]. 北京: 电子工业出版社, 2016.

[2] Zou Y B, Zhou W L, Chen X Z. Research of laser vision seam detection and tracking system based on depth hierarchical feature [J]. Chinese Journal of Lasers, 2017, 44(4): 0402009.
邹焱飏, 周卫林, 陈向志. 基于深度分层特征的激光视觉焊缝检测与跟踪系统研究[J]. 中国激光, 2017, 44(4): 0402009.

[3] Qu L, Wang K R, Chen L L, et al. Fast road detection based on RGBD images and convolutional neural network [J]. Acta Optica Sinica, 2017, 37(10): 1010003.
曲磊, 王康如, 陈利利, 等. 基于 RGBD 图像和卷积

神经网络的快速道路检测[J]. 光学学报, 2017, 37(10): 1010003.

[4] Gao H, Xue L Y. Back propagation neural network based on improved genetic algorithm fitting LED spectral model [J]. Laser & Optoelectronics Progress, 2017, 54(7): 072302.
高航, 薛凌云. 基于改进遗传算法的反向传播神经网络拟合 LED 光谱模型[J]. 激光与光电子学进展, 2017, 54(7): 072302.

[5] Xie Y. Deep learning & computer vision [M]. Beijing: China Machine Press, 2017.
叶韵. 深度学习与计算机视觉[M]. 北京: 机械工业出版社, 2017.

[6] Su M H. Research on loss function in machine learning[D]. Xi'an: Northwest University, 2015.
苏美红. 机器学习中损失函数问题研究[D]. 西安: 西北大学, 2015.

[7] Yuan L G, Xu Y, Hao Y Z. Image classification algorithm based on depth layer model [J]. Control

- Engineering of China, 2018, 25(10): 1882-1886.
- 原立格, 徐音, 郝洋洲. 基于深度层次模型的图像分类算法[J]. 控制工程, 2018, 25(10): 1882-1886.
- [8] Ren J J, Wang N. Research on cost function in artificial neural network[J]. Journal of Gansu Normal Colleges, 2018, 23(2): 61-63.
- 任进军, 王宁. 人工神经网络中损失函数的研究[J]. 甘肃高师学报, 2018, 23(2): 61-63.
- [9] Do T T, Cai M, Pham T, *et al.* Deep-6Dpose: recovering 6D object pose from a single RGB image [J/OL]. (2018-02-28)[2019-03-03]. <https://arxiv.org/abs/1802.10367>.
- [10] Wang Y M, Tan X, Yang Y, *et al.* Improving annotation for 3D pose dataset of fine-grained object categories [J/OL]. (2018-10-19) [2019-03-03]. <https://arxiv.org/abs/1810.09263>.
- [11] Li Y, Wang G, Ji X Y, *et al.* DeepIM: deep iterative matching for 6D pose estimation[C] // The European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. New York: IEEE, 2018: 683-698.
- [12] Jafari O H, Mustikovela S K, Pertsch K, *et al.* iPose: instance-aware 6D pose estimation of partly occluded objects[J/OL]. (2018-01-18)[2019-03-03]. <https://arxiv.org/abs/1712.01924>.
- [13] Krull A, Brachmann E, Nowozin S, *et al.* PoseAgent: budget-constrained 6D object pose estimation via reinforcement learning [J/OL] (2017-04-11) [2019-03-03]. <https://arxiv.org/abs/1612.03779>.
- [14] Cai P F, Duan C W. Image inpainting method based on optimal guidance rule coupled with distance constraint[J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(10): 119-125.
- 蔡鹏飞, 段朝伟. 基于最优导向法则与距离约束的图像修复算法[J]. 电子测量与仪器学报, 2018, 32(10): 119-125.
- [15] Long X, Su H S, Liu G H, *et al.* A face recognition algorithm based on angular distance loss function and convolutional neural network [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121505.
- 龙鑫, 苏寒松, 刘高华, 等. 一种基于角度距离损失函数和卷积神经网络的人脸识别算法[J]. 激光与光电子学进展, 2018, 55(12): 121505.
- [16] Zhou Y. Partial label learning algorithms based on metric learning and max-loss function [D]. Dalian: Dalian University of Technology, 2017.
- 周瑜. 基于度量学习和最大值损失函数的偏标记学习算法研究[D]. 大连: 大连理工大学, 2017.
- [17] Zhou F Y, Jin L P, Dong J. Review of convolutional neural network [J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.
- 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [18] Tekin B, Sinha S N, Fua P. Real-time seamless single shot 6D object pose prediction [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 292-301.
- [19] Rad M, Lepetit V. BB8: a scalable, accurate, robust to partial occlusion method for predicting the 3D poses of challenging objects without using depth[C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 3848-3856.
- [20] Tang M, Djelouah A, Perazzi F, *et al.* Normalized cut loss for weakly-supervised CNN segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 1818-1827.
- [21] Alhaja H A, Mustikovela S K, Geiger A, *et al.* Geometric image synthesis [J/OL]. (2018-12-01) [2019-03-03]. <https://arxiv.org/abs/1809.04696>.
- [22] Suzuki T, Kataoka H, Aoki Y, *et al.* Anticipating traffic accidents with adaptive loss and large-scale incident DB[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 3521-3529.
- [23] Kehl W, Manhardt F, Tombari F, *et al.* SSD-6D: making RGB-based 3D detection and 6D pose estimation great again[C] // Proceedings of the IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 1521-1529.
- [24] Tekin B, Sinha S N, Fua P. Real-time seamless single shot 6D object pose prediction [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 292-301.