

文章编号:1007-5321(2021)05-0101-06

DOI:10.13190/j.jbupt.2020-255

# 一种多视角高精度图片的深度估计方法

李 剑, 陈宇航

(北京邮电大学 人工智能学院, 北京 100876)

**摘要:** 针对多视图的重建中高精度图片难以有效重建的问题,提出了基于学习的深度估计方法. 该方法利用空洞卷积神经网络对图片进行特征提取,利用长短期记忆网络构建并优化三维代价体,并且采取有监督和无监督2种方式进行训练. 在2个真实场景中的多视角图片数据集上的实验结果表明,相比于传统方法和其他基于学习的方法,该网络所需的显存大大减少,因此能用于高精度图片的重建,同时,提高了模型深度预测的准确性和完整性.

**关键词:** 多视图重建; 循环神经网络; 高精度图片

**中图分类号:** TP391      **文献标志码:** A

## A Depth Estimation Method for Multi View and High Precision Images

LI Jian, CHEN Yu-hang

(School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** High-precision images are challenging to reconstruct effectively in dense multi view reconstruction. To solve the problem, a learning-based depth estimation method is proposed. In the method, the dilated convolution neural network is used to extract image features, and the long short-term memory network is applied to construct and optimize the cost volume. Besides, the supervised and unsupervised training methods are adopted. Experimental results on two real scene multi view image datasets show that the proposed method not only outperforms state-of-the-arts methods, but also is several times less in GPU memory application compared with traditional methods and other learning-based methods. Therefore, the proposed method can reconstruct high-precision images, while improving the accuracy and integrity of model depth prediction.

**Key words:** multi view reconstruction; recursive neural network; high-precision images

利用多视角图片估计物体的深度以及三维表示,是计算机视觉研究的核心问题之一. 传统方法<sup>[1]</sup>利用人工定义的相似性度量、工程正则化等流程计算二维与三维点的对应信息,进而恢复深度信息或三维点信息. 这种方法在理想的情况下获得了良好的效果,但是有其局限性,比如在弱纹理、高光和反射区域使得特征点的提取和匹配都变得非常困难,从而导致深度信息估计得不完全. 最近提出的

多视角重建算法<sup>[2-5]</sup>在像素深度预测的准确度上有了较大的提升,但是深度图的完整性依然不高,尤其是输入的图片精度比较高的情况下.

Yao等<sup>[6]</sup>提出了一种神经网络来估计图片深度,在深度图重建的完整性和质量上都得到了较大的提升,并且能在弱纹理的条件下获得较好的结果. 基于学习算法的多视角匹配方法的主要问题是在三维代价体的正则化上,此前多数网络通过多尺度三

收稿日期: 2020-12-01

基金项目: 国家自然科学基金项目(U1636106); 北京市自然科学基金项目(4182006)

作者简介: 李 剑(1976—), 男, 教授, 博士生导师, E-mail: lijian@bupt.edu.cn.

维卷积去正则化三维代价体,这种方法会消耗巨大的内存空间. 立体视觉网络 (MVSNet, multi-view stereo network)<sup>[6]</sup>在参考相机上构建三维代价体,将重建过程解耦成单视图深度估计. 但是当重建的物体图片精度变大以后, MVSNet 会失败或者耗费很长时间. 之后出现了很多 MVSNet 的相关工作, 例如循环立体视觉网络 (R-MVSNet, recurrent multi-view stereo network)<sup>[7]</sup>、点立体视觉网络 (P-MVSNet, point multi-view stereo network)<sup>[8]</sup>等.

所提算法在 MVSNet 基础上, 加入时序网络模块来提高深度预测的准确度. 视频流有严格的时间先后顺序, MVSNet 利用了多视角的信息, 但没有利用完整的时序信息, 而循环神经网络能够有效地利用时序信息, 并且在光照变化强烈、弱纹理的场景下提供更可靠的推理. 除此之外, 利用时序模块能使代价体正则化过程中需要的内存减少, 使高精度图片的训练成为可能. 另外, 某些场景下的图片序列的标签往往很难获得, 如医疗场景中利用胃镜图像重建胃模型. 基于这个原因, 所提算法在预训练模型的基础上, 利用运动恢复结构 (SFM, structure from motion) 对模型进行自监督训练, 以提高模型在不同场景下的鲁棒性.

除上述方法以外, 所提算法用深度图作为立体视觉 (MVS, multi-view stereo) 的输出. 主要原因是用深度图作为 MVS 的输出, 相比于点云数据和体素数据, 需要的资源更少, 速度更快; 其次, 利用深度图可以将像素点的深度信息很容易地转化为其他格式的数据格式, 如点云.

所提算法首先在 DTU 数据集<sup>[9]</sup>上测试, 其在深度图的完整性上超过了传统的 MVS 方法, 和 MVSNet 相近. 接下来在 Tanks and Temples 数据集<sup>[10]</sup>上利用自监督训练调整所提模型. 在该数据集中, 所提算法在效率准确率方面, 相比于传统的 MVS 方法和基于学习的 MVSNet 都有提升.

## 1 相关研究

最近卷积神经网络在计算机视觉领域中取得的成功重新激发了人们对深度图重建的兴趣. 基于学习的方法可以引入全局语义信息, 比如高光、反射等的先验信息, 以实现更具鲁棒性的匹配. 在双目立体匹配方面, 已经有了一些尝试, 利用神经网络学习得来的信息, 代替人工定义的相似度量<sup>[6-9]</sup>或正则化方法<sup>[10-12]</sup>, 都获得了很好的效果, 并且逐渐超越传统的

双目匹配方法<sup>[13-14]</sup>. 事实上, 基于卷积神经网络的方法非常适合用于双目立体匹配任务, 因为图像是预先校正过的, 所以任务就变成了水平像素的视差估计, 而不用担心相机参数的问题.

但是, 直接将双目立体匹配应用到多角度场景并不容易, 尽管可以简单地对所有图像进行预校正, 然后成对地进行双目立体匹配, 并最终合并到一个全局点上, 但是这种方法不能充分地利用多视角信息, 导致结果不准确. 与双目立体匹配不同, 多角度匹配输入的图像可以是任意位置的相机, 这给利用学习来获得三维信息的方法带来了棘手的问题. 有一些学者已经发现了这个问题, 并且尝试将卷积神经网络应用到多视角匹配中. Ji 等<sup>[13]</sup>构建了彩色体素立方体, 将所有图像像素的颜色和相机信息组合在一起, 作为网络的输入. Zhu 等<sup>[15]</sup>利用可微投影来进行端到端的训练. 但是这 2 种方法都利用规则网格表示体积, 由于三维卷积运算会消耗巨大的内存, 他们的网络拓展性并不是很好, 仅仅能处理小体积、低分辨率的对象. 在一般场景中, 表现仍然不如传统的多视角匹配算法<sup>[2-5]</sup>. 并且, 利用基于学习方法的三维重建方案会得到大量的离群点, 目前基于几何约束的方法无法有效地剔除大多数的离群点, 使深度图的边界信息和背景信息十分模糊.

## 2 方法介绍

### 2.1 模型框架

所提算法是对 MVSNet 进行一系列的拓展. 在第一阶段中, 输入图片, 利用空洞卷积模块提取图片特征, 保证图片精度不变的情况下提取不同尺度的信息. 在第二阶段中, 利用单应投影变换构造代价体. 在第三阶段中, 利用长短期记忆网络 (LSTM, long short term memory network) 模块对代价体进行正则化, 经过 softmax 模块和均值模块后生成深度图. 在整体训练中, 先在 DTU 数据集上进行有监督训练, 然后在 Tanks and Temples 数据集上利用 SFM 生成的稀疏深度图进行自监督训练. 整体流程如图 1 所示.

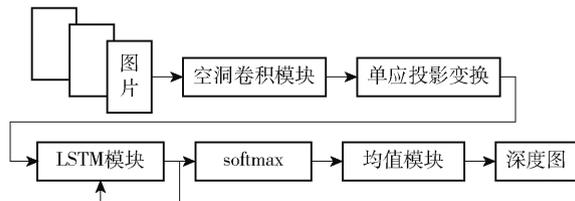


图 1 深度预测模型框架

## 2.2 算法描述

输入  $N$  个视角的有序图片, 其中  $I_{i=0}$  作为目标图片,  $I_{i=1,2,\dots,N-1}$  作为源图片. 先利用 2 个步长为 1 的卷积层简单提取图片的特征, 然后利用三层空洞卷积网络去提取不同尺度的语义信息, 空洞卷积网络的膨胀系数依次为 2, 3, 4, 且在这个过程中不会使图片精度下降. 最后得到的特征图  $F_i$  的维度为  $U \times H \times W$ , 其中,  $U$  为特征图的数量,  $H$  和  $W$  分别为输入图片的高和宽.

根据参考相机的单应矩阵, 利用平面扫描算法构造三维特征体,  $N$  张特征图可以形成  $N$  个特征体  $\{V_i\}_{i=1}^N$ , 利用特征体在深度方向上根据方差累加到代价体  $C$  中, 计算公式为

$$C = \frac{\sum_{i=1}^N (V_i - \bar{V})^2}{N} \quad (1)$$

所提算法将利用改进后的 LSTM 模块<sup>[15]</sup> 对生成的代价体  $C$  进行正则化去除噪声, 整体流程如图 2 所示.

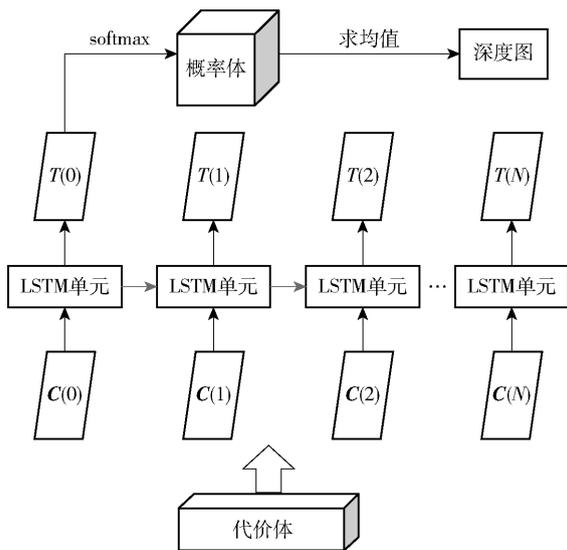


图 2 正则化模块框架

首先定义接下来要使用的符号.  $D$  为深度先验,  $\{T(i)\}_{i=0}^{D-1}$  为第  $i$  步的输出,  $T(i)$  不仅仅依赖于当前帧, 也依赖于此前的状态  $T(0, \dots, i-1)$ .  $I(i)$ ,  $F(i)$ ,  $O(i)$  分别对应 LSTM 的输入门、遗忘门和输出门. ‘ $\odot$ ’, ‘ $[\ ]$ ’ 和 ‘ $*$ ’ 分别为点积、级联和矩阵乘法.

输入门的作用是从当前帧  $\hat{T}(i)$  中选取有效数据传递给当前的单元状态  $T(i)$ :

$$I(i) = \sigma(W_I[T(i), T(i-1)] + B_I) \quad (2)$$

$$\hat{T}(i) = \tanh(W_T[T(i), T(i-1)] + B_T) \quad (3)$$

其中:  $\sigma$  为 sigmoid 激活函数,  $W_I$  为输入门中的权重,  $B_I$  为输入门中的偏移量,  $W_T$  为当前状态单元  $T(i)$  的权值,  $B_T$  为当前状态单元  $T(i)$  的偏移量.

遗忘门的作用是从之前的状态单元  $T(i-1)$  中筛选出一些无用的信息, 然后结合从经过输入门的信息, 生成新当前的单元状态  $T(i)$ :

$$F(i) = \sigma(W_F[T(i), T(i-1)] + B_F) \quad (4)$$

$$T(i) = F(i) \odot T(i-1) + I(i) \odot \hat{T}(i) \quad (5)$$

其中:  $W_F$  为遗忘门的权重,  $B_F$  为遗忘门的偏移量.

输出门的作用是控制从当前状态单元  $T(i)$  输出的信息, 最终的输出为

$$O(i) = \sigma(W_O[T(i), T(i-1)] + B_O) \quad (6)$$

$$O(i) = \sigma(W_O[T(i), T(i-1)] + B_O) \quad (7)$$

其中:  $W_O$  为输出门的权重,  $B_O$  为输出门的偏移量.

在所提算法中, 利用 3 个不同尺度的 LSTM 模块获得不同尺度下的信息, 利用 2 个 LSTM 模块收集多尺度信息, 从而提高了预测的鲁棒性和准确性. 并且, 由于将代价体拆成深度上不同的二维代价体, 有效地减少了内存的使用, 使更大精度图片的训练成为可能. 最终获得正则化过后的三维代价体  $\{T(i)\}_{i=0}^{D-1}$ .

将正则化过后的代价体  $\{T(i)\}_{i=0}^{D-1}$  通过 softmax 层, 生成相应的概率体  $P$ . 利用在深度方向上的期望值作为该像素的深度估计值来生成深度图  $D$ :

$$D = \sum_{d=d_{\min}}^{d_{\max}} d \times P(d) \quad (8)$$

其中  $d_{\min}$  和  $d_{\max}$  分别为先验深度的最小值和最大值.

最终使用交叉熵作为模型的损失函数:

$$L = - \sum_{x \in x_v} \sum_{i=0}^{D-1} G(i, x) \lg P(i, x) \quad (9)$$

其中:  $x_v$  为真值中有效点的集合,  $G(i, x)$  为热编码后的深度  $i$  在像素  $x$  上的真值,  $P(i, x)$  为三维概率体的深度  $i$  在像素  $x$  上的估计值.

## 3 数据集

目前的多视角重构任务数据集的三维信息为点云和网络格式. 需要把这些三维信息转为深度信息. DTU 数据集是一个室内场景 MVS 数据集, 包含了超过 100 个场景的不同光照强度的数据. 由于图形处理器 (GPU, graphics processing unit) 的限制, 实验中把 DTU 数据集中图片的大小降至原图的 1/16 进行训练, 原图大小为  $W \times H = 640 \times 512$ , 处理后的

大小为  $W \times H = 160 \times 128$ . DTU 数据集提供的标签为点云格式,利用泊松表面先重建生成网状表面,然后再将其映射到不同视角下,生成深度图,作为训练数据.在泊松重建阶段,参数最大深度为 11,修剪因子为 9.5.

深度范围的选择根据上述过程生成的深度图来设定,根据 DTU 数据集的结果,设置的先验深度范围  $[d_{\min}, d_{\max}]$  为  $[425 \text{ mm}, 745 \text{ mm}]$ ,精度为 2.5 mm,深度值为 128.

另外一部分需要预处理的数据为部分 Tanks and Temples 数据集中的数据,这是一个室外场景数据集,包含了更加复杂的场景和光照情况,和真实场景非常接近,测试的数据集场景包括雕塑、坦克、马、操场、火车、教堂等.处理后的图片大小为  $W \times H = 160 \times 128$ . 预处理后的数据集如表 1 所示.

表 1 数据集中图片数量统计

数据集	训练集	验证集	测试集	总数
DTU	34 000	4 266	4 266	42 532
Tanks and Temples	10 000	/	2 000	12 000

## 4 实验结果与分析

### 4.1 实验设置

在 DTU 数据集上进行监督训练得到有监督模型,在 Tanks and Temples 数据集上基于有监督模型进行自监督训练调整得到无监督模型.输入图片的大小  $W \times H = 160 \times 128$ ,输出的深度图大小  $W \times H = 160 \times 128$ ,每次输入的图片数量  $N = 3$ ,包括 1 张目标图片和 2 张源图片.所提算法使用 Adam 作为优化器进行训练,学习率为 0.001,共进行 8 次迭代训练.所提算法在 DTU 数据集和 Tanks and Temples 数据集上进行测试,评价指标使用准确率和召回率.准确率为深度预测正确的像素点和所有预测出的像素点总数之比,召回率为深度预计正确的像素点和真值的像素点总数之比.允许接受的误差阈值分别设置为 1 mm 和 2 mm.

### 4.2 实验结果及分析

为了验证所提算法的有效性,一共设置了 3 组实验来验证所提算法的性能.

#### 4.2.1 有监督训练实验

第 1 组实验为有监督训练算法模型在 DTU 数据集和 Tanks and Temples 数据集上的表现,对比算法为传统方法的 Gipuma<sup>[4]</sup>, Tola<sup>[17]</sup>, Colmap<sup>[5]</sup> 算法

和其他 2 种基于学习的 MVSNet<sup>[6]</sup> 算法和 R-MVSNet<sup>[7]</sup> 算法.

表 2 所示为不同的算法在 DTU 数据集上的性能对比.由表 2 可知,在 DTU 数据集上,利用传统方法 Gipuma、Tola 和 Colmap 生成的深度图,在准确度上要好于基于学习的三维重建方法,但是在召回率上表现不如后者.这意味着,基于学习的三维重建方法能找到更多的点,因此深度图的完整性更高.相比于 MVSNet 算法和 R-MVSNet 算法,所提算法在准确度上和这 2 种算法相差不大,但是在召回率上表现更好,提升了约 5%,因此在保证准确度的前提下,可以生成更加完整的深度图.

表 2 所提算法和对比算法在 DTU 数据集上的性能对比 %

算法	阈值 = 1 mm		阈值 = 2 mm	
	准确率	召回率	准确率	召回率
Gipuma	94.65	70.64	96.42	74.16
Tola	90.49	68.07	93.94	73.61
Colmap	86.31	65.34	89.62	72.27
MVSNet	86.46	75.69	91.06	80.25
R-MVSNet	87.23	79.04	94.29	82.92
所提算法	86.72	82.99	93.52	84.49

表 3 所示为不同的算法在 Tanks and Temples 数据集上的性能对比,其中误差阈值设置为 1 mm.由表 2 和表 3 的比较可知,与在 DTU 数据集上的性能对比,在 Tanks and Temples 数据集中所有算法的性能指标都大大下降,准确率平均下降约为 15%,召回率下降约为 30%.但在这类陌生场景中,基于学习的三维重建方法在准确率和召回率上整体优于传统方法.所提算法在准确率上高于 MVSNet 算法大约 3%,略低于 R-MVSNet 算法,在召回率的表现上高于另外 2 种算法超过 4%,因此,在陌生场景中,

表 3 所提算法和对比算法在 Tanks and Temples 数据集上的性能对比 %

算法	准确率	召回率
Gipuma	76.21	43.44
Tola	74.32	43.48
Colmap	70.84	42.14
MVSNet	74.40	43.76
R-MVSNet	78.39	48.40
所提算法	77.45	52.19

所提算法能在保持较好准确度的同时,预测更加完整的深度图。

#### 4.2.2 自监督训练实验

第 2 组实验为有监督训练算法与自监督训练算法在 Tanks and Temples 数据集上的表现比较,误差阈值设置为 1 mm。

表 4 所示为所提算法在有监督训练和无监督训练情况下的对比结果。由表 4 可知,在陌生场景中对网络进行自监督训练调优,能够提高所提算法对深度预测的准确度和完整性,准确率从 77% 提升到了 80%,而召回率从 52% 提升到了 62%,在准确率提高的同时,像素点的完整性也得到了较大的提升,使深度图的预测更为完整,所以利用自监督训练对模型进行调优是有效的。

**表 4 Tanks and Temples 数据集上有监督和自监督算法结果** %

算法	准确率	召回率
有监督算法	77.45	52.19
自监督算法	80.39	62.48

#### 4.2.3 运行时间及内存消耗实验

第 3 组实验为有监督训练模型在 DTU 数据集上测试时运行时间和 GPU 内存消耗量的表现,对比算法为基于学习的 MVSNet 算法和 R-MVSNet 算法,测试的图片大小为 1 600 × 1 200。

表 5 所示为不同算法输入一张 DTU 数据集中的图片时的运行时间和 GPU 内存消耗结果。由表 5 可知,在保证准确率的情况下,所提算法在内存消耗量和运行时间上,要远远小于 MVSNet,在运行时间上略慢于 R-MVSNet,但是在内存消耗上,则比 R-MVSNet 要少。

**表 5 所提算法和对比算法在 DTU 数据集上的运行时间和 GPU 内存消耗对比**

算法	运行时间/s	GPU 内存消耗/GB
MVSNet	29.15	15.4
R-MVSNet	2.1	6.7
所提算法	2.6	5.2

综上所述,利用所提算法能够在陌生场景中保证准确率和召回率的前提下,大大加快模型的运行速度并且减少模型所需的 GPU 内存。并且,加入自监督训练能够在陌生场景中有效地提高预测的准确度和完整性。

## 5 结束语

所提算法利用卷积空洞神经网络提取不同尺度下的特征,利用改进后的时序模块 LSTM 顺序地处理三维代价体,使得模型能够获取上下文的信息,提高模型的准确度的同时,大大降低了模型运行时的 GPU 内存消耗。此外利用 SFM 进行自监督训练能够有效地在陌生场景中对所提算法进行调优,提高模型的鲁棒性。在室内数据集 DTU 和室外数据集 Tanks and Temples 上的实验,证明了所提算法的有效性,在提高了准确度和完整性的同时,降低了对 GPU 内存的消耗。未来的工作中,将更多地探索不同的自监督的方法,进行不同的实验,一方面尝试在自监督过程中加入光流、几何等约束条件,另一方面尝试利用深度学习方法在自监督过程中加入不同的特征,提高自监督的效果和泛化能力。

### 参考文献:

- [1] Hirschmuller H. Stereo processing by semiglobal matching and mutual information [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 30(2): 328-341.
- [2] Furukawa Y, Ponce J. Accurate, dense, and robust multi-view stereopsis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 32(8): 1362-1376.
- [3] Vu H-H, Labatut P, Pons J-P, et al. High accuracy and visibility-consistent dense multiview stereo [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(5): 889-901.
- [4] Galliani S, Lasinger K, Schindler K. Massively parallel multiview stereopsis by surface normal diffusion [C] // IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 873-881.
- [5] Schönberger J L, Zheng Enliang, Frahm J M, et al. Pixel-wise view selection for unstructured multi-view stereo [C] // European Conference on Computer Vision. Amsterdam: Springer, 2016: 501-518.
- [6] Yao Yao, Luo Zixin, Li Shiwei, et al. MVSnet: depth inference for unstructured multi-view stereo [C] // European Conference on Computer Vision. Munich: Springer, 2018: 767-783.
- [7] Yao Yao, Luo Zixin, Li Shiwei, et al. Recurrent MVSnet for high-resolution multi-view stereo depth inference [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 5525-5534.

- [8] Luo Keyang, Guan Tao, Ju Lili, et al. P-MVSnet: learning patch-wise matching confidence aggregation for multi-view stereo[C]//IEEE/CVF International Conference on Computer Vision. Long Beach: IEEE, 2019: 10452-10461.
- [9] Aanæs H, Jensen R R, Vogiatzis G, et al. Large-scale data for multiple-view stereopsis[J]. International Journal of Computer Vision, 2016, 120(2): 153-168.
- [10] Knapitsch A, Park J, Zhou Qianyi, et al. Tanks and temples: benchmarking large-scale scene reconstruction[J]. ACM Transactions on Graphics, 2017, 36(4): 1-13.
- [11] Angelova A, Long P M. Benchmarking large-scale fine-grained categorization[C]//IEEE Winter Conference on Applications of Computer Vision. Steamboat Springs: IEEE, 2014: 532-539.
- [12] Hartmann W, Galliani S, Havlena M, et al. Learned multi-patch similarity[C]//IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 1586-1594.
- [13] Ji Mengqi, Gall J, Zheng Haitian, et al. Surfnet: an end-to-end 3D neural network for multiview stereopsis [C] // IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2307-2315.
- [14] Choi S, Kim S, Park K, et al. Learning descriptor, confidence, and depth estimation in multi-view stereo [C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018: 276-282.
- [15] Zhu Guangming, Zhang Liang, Shen Peiyi, et al. Multimodal gesture recognition using 3-D convolution and convolutional LSTM[J]. IEEE Access, 2017, 5: 4517-4524.
- [16] Konolige K, Agrawal M. FrameSLAM: from bundle adjustment to real-time visual mapping [J]. IEEE Transactions on Robotics, 2008, 24(5): 1066-1077.
- [17] Tola E, Strecha C, Fua P. Efficient large-scale multi-view stereo for ultra high-resolution image sets[J]. Machine Vision and Applications, 2012, 23(5): 903-920.

(上接第 87 页)

- [7] Grover A, Leskovec J. Node2vec: scalable feature learning for networks[C]//22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016: 855-864.
- [8] Liu Hao, Tong Yongxin, Zhang Panpan, et al. Hydra: a personalized and context-aware multi-modal transportation recommendation system[C]//25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Anchorage: ACM, 2019: 2314-2324.
- [9] Sun Hanxiao, Wang Wenjie, Shi Zhongzhi. Parallel factorization machine recommended algorithm based on MapReduce[C]//10th International Conference on Semantics, Knowledge and Grids. Beijing: IEEE, 2014: 120-123.
- [10] 曹燕,董一鸿,邬少清,等. 动态网络表示学习研究进展[J]. 电子学报, 2020, 48(10): 2047-2059.  
Cao Yan, Dong Yihong, Wu Shaoqing, et al. Dynamic network representation learning: a review [J]. Acta Electronica Sinica, 2020, 48(10): 2047-2059.
- [11] Mikolov T, Sutskever I, Chen Kai, et al. Distributed representations of words and phrases and their compositionality[C]//26th International Conference on Neural Information Processing Systems. North Miami Beach: Curran Associates Incorporated, 2013: 3111-3119.
- [12] Cui Peng, Wang Xiao, Pei Jian, et al. A survey on network embedding[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(5): 833-852.
- [13] Ke Guolin, Meng Qi, Finley T, et al. LightGBM: A highly efficient gradient boosting decision tree [C] // 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Incorporated, 2017: 3146-3157.
- [14] Chen Tianqi, Guestrin C. XGBoost: A scalable tree boosting system[C]//22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016: 785-794.
- [15] Yang Gun, Zhou Fangrong, Ma Yi, et al. Identifying lightning channel-base current function parameters by Powell particle swarm optimization method [J]. IEEE Transactions on Electromagnetic Compatibility, 2018, 60(1): 182-187.
- [16] Meier L, Geer S V D, Bhlmann P, et al. The group lasso for logistic regression[J]. Journal of the Royal Statistical Society Series B (Statistical Methodology), 2008, 70(1): 53-71.
- [17] Lee J S. AUC4. 5: auc-based C4. 5 decision tree algorithm for imbalanced data classification [J]. IEEE Access, 2019: 106034-106042.
- [18] Breiman L. Random forest [J]. Machine Learning, 2001, 45: 5-32.