

## 基于卷积神经网络和流形排序的图像检索算法

刘 兵<sup>1,2</sup>, 张 鸿<sup>1,2\*</sup>

(1. 武汉科技大学 计算机科学与技术学院, 武汉 430065;

2. 武汉科技大学 智能信息处理与实时工业系统湖北省重点实验室, 武汉 430065)

(\* 通信作者电子邮箱 46476522@qq.com)

**摘 要:** 针对基于内容的图像检索 (CBIR) 中低层视觉特征与用户对图像理解的高层语义不一致以及传统的距离度量方式难以真实反映图像之间相似程度等问题, 提出了一种基于卷积神经网络 (CNN) 和流形排序的图像检索算法。首先, 将图像输入 CNN, 通过多层神经网络对图像的监督学习, 提取网络中全连接层的图像特征; 其次, 对图像特征进行归一化处理, 然后用高效流形排序 (EMR) 算法对查询图像所返回的结果进行排序; 最后, 根据排序的结果返回最相似的图像。在 corel 数据集上, 深度图像特征比基于场景描述的图像特征的平均查准率 (mAP) 提高了 53.74%, 流形排序比余弦距离度量方式的 mAP 提高了 18.34%。实验结果表明, 所提算法能够有效地提高图像检索的准确率。

**关键词:** 图像检索; 深度学习; 卷积神经网络; 特征提取; 流形排序

**中图分类号:** TP391.413 **文献标志码:** A

## Image retrieval algorithm based on convolutional neural network and manifold ranking

LIU Bing<sup>1,2</sup>, ZHANG Hong<sup>1,2\*</sup>

(1. School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan Hubei 430065, China;

2. Hubei Province Key Laboratory of Intelligent Information Processing and Real-Time Industrial System,  
Wuhan University of Science and Technology, Wuhan Hubei 430065, China)

**Abstract:** In Content-Based Image Retrieval (CBIR), the low-level visual features are not consistent with the high-level semantic features captured by human, and it is difficult to reflect the similarity of images by traditional distance measurements. To solve these problems, an image retrieval algorithm based on Convolutional Neural Network (CNN) and manifold ranking was proposed. Firstly, the image dataset was put into CNN, image features were extracted through the fully connected layers of the network after supervised learning; secondly, the image features were normalized and then Efficient Manifold Ranking (EMR) algorithm was used to return the ranked scores for query images; finally, the most similar images were returned to users according to the scores. In corel dataset, the mean Average Precision (mAP) of deep image feature was 53.74% higher than that of the scene descriptor features, and the mAP of efficient manifold ranking was 18.34% higher than that of the cosine distance. The experimental results show that the proposed algorithm can effectively improve the accuracy of image retrieval.

**Key words:** image retrieval; deep learning; Convolutional Neural Network (CNN); feature extraction; manifold ranking

## 0 引言

基于内容的图像检索 (Content-Based Image Retrieval, CBIR) 是利用图像的低层语义特征, 如颜色、形状、纹理等对图像进行检索的技术。近年来, 众多研究者在基于内容的图像检索中取得了大量的研究成果。由于用户高层语义和图像的低层特征表示之间的差距, 即“语义鸿沟” (semantic gap), 导致实际检索效果与理想结果差距较大。图像特征提取是图像检索中关键的一步, 好的图像特征表示能够提升检索效果。最近较好的局部图像特征表示, 如词袋模型 (Bag-of-Words, BoW)<sup>[1-2]</sup>, 用的是尺度不变特征转换 (Scale-Invariant Feature Transform, SIFT)<sup>[3]</sup> 或加速稳健特征 (Speeded Up Robust Feature, SURF)<sup>[4]</sup> 等局部图像特征。此外, 特征之间的相似性度量也是图像检索中重要的一步。

深度学习<sup>[5]</sup> 是一种利用多个隐层的深度神经网络完成

学习任务的方法, 是机器学习研究中的一个新的领域, 其目的在于建立模拟人的大脑进行学习的神经网络, 模仿人脑的机制来解释数据, 例如图像、声音和文本。为此, 深度学习有可能成为跨越“语义鸿沟”的桥梁<sup>[6-7]</sup>。深度学习在图像分类<sup>[8]</sup> 等问题中取得了比较好的效果。图像分类和图像检索是两个相对独立又存在着某种关联的问题。本文提出了一种基于卷积神经网络 (Convolutional Neural Network, CNN) 模型来提取图像特征的方法。由于图像之间的相似性程度一般是采用向量之间的距离来计算的, 常用的度量方法有余弦距离和欧氏距离等。对于图像而言, 简单的向量之间的距离难以真实地反映图像之间的相似程度, 为此, 本文在深度图像特征的基础上采用基于流形学习的相似性度量方法, 即流形排序 (Manifold Ranking, MR)。流形排序能够更真实地反映图像之间的相似程度。实验表明, 深度图像特征和流形排序方法能在一定程度上提高图像检索的精度。

收稿日期: 2015-07-24; 修回日期: 2015-09-12。 **基金项目:** 国家自然科学基金资助项目 (61003127, 61373109)。

**作者简介:** 刘兵 (1992 -), 男, 湖北广水人, 硕士研究生, 主要研究方向: 深度学习、图像处理; 张鸿 (1979 -), 女, 湖北襄阳人, 教授, 博士, CCF 会员, 主要研究方向: 跨媒体检索、机器学习、数据挖掘。

## 1 基于卷积神经网络的特征分析

深度学习是机器学习领域中的一系列算法,主要是基于多层的人工神经网络。本文使用的是深度学习中的卷积神经网络,它是一种前馈神经网络,主要由多个卷积层和全连接层组成。同一层中的某些神经元之间的连接的权重是共享的。一个前馈神经网络可以被认为是一系列函数的组合,定义如下:

$$f(\mathbf{x}) = f_{\ell}(\cdots f_2(f_1(\mathbf{x}; \mathbf{w}_1); \mathbf{w}_2) \cdots; \mathbf{w}_{\ell}) \quad (1)$$

每一个函数 $f_{\ell}$ 都有一个基准输入 $\mathbf{x}_{\ell}$ 和一个参数向量 $\mathbf{w}_{\ell}$ ,其输出为 $\mathbf{x}_{\ell+1}$ ,参数 $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{\ell})$ 是从目标问题中学习得来的。卷积神经网络中的数据和函数的结构略有不同,其中数据 $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ 通常为图像、声音等。每个 $\mathbf{x}_i$ 都是 $M \times N \times K$ 的矩阵,像素为 $M \times N$ ,通道数为 $K$ 。设向量 $\mathbf{y}$ 为卷积输出,过滤器是三维的,特征图(feature map)是 $K$ 通道的,若有 $K'$ 个过滤器,产生 $K'$ 维的特征图 $\mathbf{y}$ ,则:

$$y_{i'j'k'} = \sum_{ijk} w_{ijk} x_{i+i', j+j', k+k'} \quad (2)$$

卷积神经网络的损失函数可以定义为:

$$L(\mathbf{w}) = \frac{1}{n} \sum_{i=1}^n \ell(\mathbf{z}_i, f(\mathbf{x}_i; \mathbf{w})) \quad (3)$$

其中 $\mathbf{Z} = f(\mathbf{x}, \mathbf{w})$ 。在卷积神经网络中,对于输入特征图 $\mathbf{x}$ ,要进行归一化操作,公式如下:

$$y_{ijk} = x_{ijk} / \left( \kappa + \alpha \sum_{k \in G(k)} x_{ijk}^2 \right)^{\beta} \quad (4)$$

其中 $G(k) = [k - \lfloor \rho/2 \rfloor, k + \lceil \rho/2 \rceil] \cap \{1, 2, \cdots, K\}$ 。

卷积神经网络中还有两个重要操作:一个是非线性激活函数(Rectified Linear Unit, ReLU);一个是池化(Pooling)。ReLU激活函数的定义如下:

$$y_{ijk} = \max(0, x_{ijk}) \quad (5)$$

池化操作是对卷积过程中提取的不同位置的图像特征进行聚集统计。池化能够降低卷积特征的维度,同时还能防止数据过拟合。文中所使用的是最大池(max Pooling)操作,其公式如下:

$$y_{ijk} = \max\{y_{i'j'k'}; i \leq i' < i+p, j \leq j' < j+p\} \quad (6)$$

本文使用的卷积神经网络基于 MatConvNet 工具箱<sup>[9]</sup>。MatConvNet 是用 Matlab 开发的一个卷积神经网络工具包,简单高效,并且能够运行和学习一些常用的卷积神经网络。图1展示了本文中卷积神经网络的构成及图像检索的算法流程。其中:C1~C5代表网络中的卷积层,FC6~FC8代表网络中的全连接层。

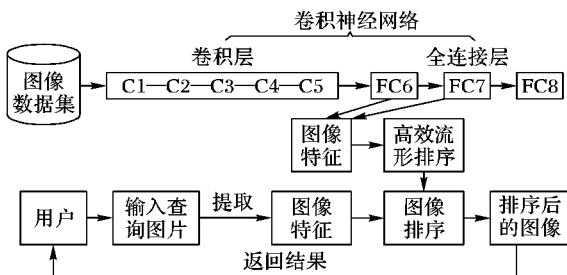


图1 基于卷积神经网络和流形排序的图像检索算法

卷积神经网络模型来自 Chatfield 等<sup>[10]</sup>的视觉几何组(Visual Geometry Group, VGG)项目。从文献[10]和图1中卷积部分示意图,可以看出卷积神经网络共有8层,分别为5个卷积层和3个全连接层。由于输入卷积神经网络的图像要求尺寸一致,故 RGB 图像进入卷积神经网络前,首先是对图

像进行预处理操作,将原始的 RGB 图像的尺寸统一调整为 $224 \times 224$ 。卷积神经网络接收图片后,把图像的每个通道当作一个特征图,分通道单独处理,然后进入卷积层进行卷积操作。前面两个卷积层每层进行卷积运算,之后用 ReLU 函数激活,然后是池化操作;接着的两个卷积层先卷积,再 ReLU 操作;最后一个卷积层在 ReLU 之后进行池化操作。最后为全连接层:前两个全连接层使用 ReLU 激活函数,最后一个全连接层使用的是 softmax 分类器。

最近的一些研究<sup>[11-12]</sup>表明,卷积神经网络可以从图像中学习到的中层语义信息,卷积神经网络中全连接层中的图像特征能够有效地用于图像检索等任务中。由于卷积神经网络提取的特征是通过监督学习而得到的,不能够确定哪一层提取的特征更有利于提升检索效果,故本文选取了第6和第7两个全连接层来提取图像的特征,其图像特征维数均为4096维,把提取的特征分别记为第6层特征(CNN\_L6)和第7层特征(CNN\_L7)。通过对这两个不同层提取的图像特征的检索效果的对比,最终来确定哪一层的检索效果更好。为了方便后续的相似性度量,对提取出来的图像特征进行 L2 归一化的处理。

## 2 基于流形排序的相似性度量

传统的图像的相似性度量方法有余弦距离、欧氏距离等,通过特征向量之间的距离来判断图像的相似程度。由于单一的特征向量之间的距离并不能够真实地反映图像之间的相似程度,故本文使用了一种基于流形距离的流形排序<sup>[13-14]</sup>方法来度量图像之间的相似度。

基于流形的排序方法是由 Zhou 等<sup>[13]</sup>首次提出来的,它通过构造一个数据间的流形结构来计算排序得分。首先构造出一个带权值的网络,并且给初始节点分配一个分数,其他需要排序节点的分数设为0;然后每个节点将自己的分数传到相邻的节点直到整个网络处于平衡的状态。除了初始节点外的其他节点按最后各自的分数排序。分数越高,排序越靠前。

设数据集的构成 $\chi = \{x_1, x_2, \cdots, x_n\} \subset \mathbf{R}^m$ ,  $x_i$ 表示初始查询,  $x_i (1 \leq i \leq n)$  为候选查询。设 $d: \chi \times \chi \rightarrow \mathbf{R}$ 表示 $\chi$ 上的一个度量,其中 $d$ 表示 $x_i$ 和 $x_j$ 之间的距离。排序函数为 $r: \chi \rightarrow \mathbf{R}$ ,对于每一个查询计算其排序得分。最后定义一个向量 $\mathbf{y} = [y_1, y_2, \cdots, y_n]^T$ ,若 $x_i$ 是初始查询则 $y_i = 1$ ;否则为0。构造一个图 $G = (V, E)$ ,其中 $V$ 表示为数据集 $\chi$ ,  $E$ 的权重表示为相似度矩阵 $\mathbf{W} = [w_{ij}]_{n \times n}$ 。计算图的相似度矩阵 $\mathbf{D} = \text{diag}\{d_{11}, d_{22}, \cdots, d_{nn}\}$ ,其中 $d_{ii} = \sum_j w_{ij}$ 。归一化图的相似度矩阵 $\mathbf{W}$ 为 $\mathbf{S} = \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}$ 。然后利用式(7)进行迭代,直到收敛为止。

$$r(t+1) = \alpha \mathbf{S} r(t) + (1-\alpha) \mathbf{y} \quad (7)$$

其中 $\alpha \in [0, 1]$ 。令 $r_i^*$ 代表数列 $\{r_i(t)\}$ 的极值,最后根据 $r_i^*$ 对数据点 $x_i$ 进行排序。

流形排序算法在实际应用中存在两个问题:1) 图的构造基本固定,不能有效处理可扩展的数据集;2) 排序函数在计算排序得分的过程中要消耗较长的时间,导致检索的速度较慢。本文采用文献[14]中的高效流形排序(Efficient Manifold Ranking, EMR)算法。设权重矩阵 $\mathbf{Z} \in \mathbf{R}^{d \times n}$ ,其中 $d$ 表示固定点的个数。令 $\mathbf{W} = \mathbf{Z}^T \mathbf{Z}$ ,如果两个数据点相关,那么至少存在一个固定点。设 $\mathbf{H} = \mathbf{Z} \mathbf{D}^{-1/2}$ ,  $\mathbf{S} = \mathbf{H}^T \mathbf{H}$ ,那么排序函数 $r$ 可表示为:

$$r^* = (\mathbf{I}_n - \alpha \mathbf{S})^{-1} \mathbf{y} \quad (8)$$



其中  $I_n$  表示是一个  $n \times n$  的单位矩阵。用式(8)来计算图像之间的排序分数。

### 3 实验与分析

#### 3.1 实验数据集

实验中选用了两种数据集: corel 和 caltech256。其中 corel 图像集总共有 10 000 张图片,共有 100 类,每类均为 100 张图片,同一类的图片含有相同的语义信息,包含的类别有花、鸟、沙滩等;caltech256 图像集总共有 30 607 张图片,共有 257 类。

#### 3.2 性能评价指标

图像检索一般采用查准率 *precision*、召回率 *recall* 和平均查准率 (mean Average Precision, mAP) 作为评价指标。

$$precision = (a/b) \times 100\%$$

$$recall = (a/c) \times 100\%$$

$$mAP(Q) = \frac{1}{|Q|} \sum_{j=1}^{|Q|} \frac{1}{m_j} \sum_{k=1}^{m_j} Precision(R_{jk})$$



图2 corel 数据集返回的 top10 图像实例



图3 caltech256 数据集返回的 top10 图像实例

#### 3.3.2 检索平均查准率

对 corel 图像数据集中所有的图像类别都进行了测试,平均查准率结果如表1所示。表1中:CNN-L6 和 CNN-L7 分别表示从卷积神经网络中的第6层和第7层提取出的图像特征;Cosine 表示余弦距离的度量方法,EMR 表示用高效流形排序算法对图像特征进行排序;gist + Cosine 和 gist + EMR 表示用基于 Gabor 变换的 gist<sup>[15]</sup> 图像特征的实验结果,gist 是一种基于场景的特征描述符。

表1 corel 数据集中检索 mAP 比较 %

图像特征和度量方法	mAP	图像特征和度量方法	mAP
CNN-L6 + Cosine	51.74	CNN-L7 + EMR	71.34
CNN-L6 + EMR	70.59	gist + Cosine	13.17
CNN-L7 + Cosine	53.00	gist + EMR	17.60

从表1中可以看出,使用余弦距离度量方式的图像检索的平均查准率在 50% 左右,使用 gist 特征的平均查准率在

其中: $a$  是返回结果中正确结果的个数, $b$  是返回结果的个数, $c$  表示系统中相关结果的总数, $Q = \{d_1, d_2, \dots, d_{m_j}\}$ ,  $R_{jk}$  表示检索到  $d_k$  元素时的排序结果。

#### 3.3 实验结果与分析

##### 3.3.1 检索实例

图2是 corel 数据集中的检索实例。从图2(b)和(c)中可以看出,在 corel 数据集中,EMR 算法的 top10 查准率要高于余弦距离,使用余弦距离检索返回的图像中有 3 幅不相关图像,而 EMR 算法返回的图像都是相关的。

图3是 caltech256 数据集中的检索实例。该查询图像包含人和枪两个元素,属于枪这一类别,返回给用户的图像应该是含有枪的图像。从图3(b)和(c)中可以看出,EMR 算法返回的图像基本都与枪相关,而余弦距离返回的图像大多数都不相关。余弦距离单一的向量之间的距离并不能较好地反映图像之间的相似程度,而 EMR 算法在一定程度上能够克服这个缺点。

13% ~ 17%,而基于深度图像特征的 EMR 算法的平均查准率在 70% 左右。对比可知,EMR 算法检索的平均查准率要高于余弦距离,它对于图像的检索效果有着明显的提升。深度图像特征是从图像数据中自动学习到的特征表示,因而比 gist 特征的检索效果要好。对比表1中相同方法不同层数的深度图像特征可知,第7层的深度图像特征的检索效果都要略优于第6层。卷积网络层数越深,网络中的神经元(参数)也就越多,也就能更好地表现语义特征。

本文对两种度量方式的 topK 查准率作了测试,topK 表示的是检索中返回排序中的前 K 幅相似图像的查准率。实验结果如图4所示,横轴 K 代表返回图像的数目。可以看出,EMR 算法的 topK 查准率比余弦距离好。当返回图片的数目增多时,EMR 算法的 topK 查准率同样优于余弦距离。但由于返回图片数目变多,检索到不相关的图片的可能性也增大,所以也导致了查准率的下降。

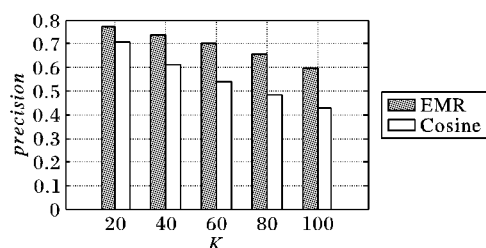


图4 corel数据集topK查准率

对于 caltech256 数据集,本文按照文献[16]的方法从257类中分别选取了10类、20类和50类进行测试,最后又对数据集中所有的类别进行了检索实验,结果如表2所示,其中OASIS(Online Algorithm for Scalable Image Similarity)表示在线可扩展图像相似算法,来自文献[16]的实验结果,使用的是BoW模型。首先,对比表中深度图像特征与gist特征及BoW模型的实验结果,可以看出深度图像特征的检索效果都比传统图像特征好。CNN提取的图像特征与传统的手工特征最大的不同在于它所得到的特征是从图像中不断学习而得到的。深度学习具有强大的学习能力和高效特征表达能力,能从原始数据到抽象的语义中逐层提取信息,在一定程度上可以减少“语义鸿沟”所带来的影响。其次,对比表2中前4行,可以看出EMR的检索效果比余弦距离的效果要好,即流形排序比传统的距离度量方式要好;第7层提取的图像特征比第6层的图像特征的检索效果好。当检索类别增多时,检索到的不相关图像的几率会相应增加,检索的查准率一般会下降。

表2 caltech256数据集中检索mAP比较

特征和度量方法	10类	20类	50类	257类
CNN-L6 + Cosine	43.80	49.55	37.88	39.43
CNN-L6 + EMR	50.60	54.93	44.26	44.18
CNN-L7 + Cosine	46.49	51.53	40.50	41.48
CNN-L7 + EMR	53.31	57.19	46.31	45.91
gist + EMR	1.38	13.58	3.78	7.14
BoW + OASIS <sup>[16]</sup>	33.00	21.00	12.00	—

### 3.3.3 召回率

两个数据集的召回率实验结果如图5、6所示,所用的图像特征是CNN-L7,而且图6中由于gist特征的召回率太低,故没有画出。从图5、6中可以看出,随着检索图像数量的增加,召回率也随之增加,基于深度图像特征的EMR算法的召回率要高于对比方法。

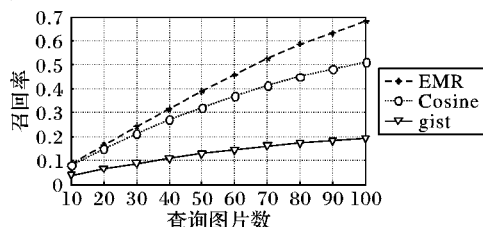


图5 corel数据集召回率

### 3.3.4 检索效率

本文使用的操作系统为Windows 8.1 64位, Matlab版本为2013b,硬件为Intel Core i3-3240 CPU 3.4 GHz, 8 GB内存,无GPU。在提取深度图像特征过程中,没有使用GPU加速,提取图像特征的速度较慢,提取一幅图像的特征耗时约为0.68 s。表3展示了两种方法在不同数据集上检索所有类别

所用的总时间。结果表明,EMR比余弦距离的速度快很多,在较大的数据集上也有较快的检索速度。

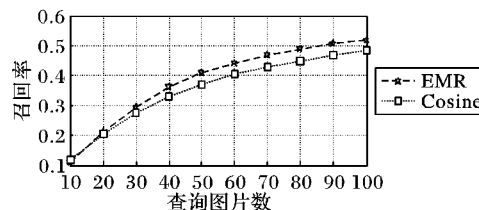


图6 caltech256数据集召回率

表3 检索时间比较

数据集	EMR	Cosine
corel	1356.35	5976.97
caltech256	4792.89	64378.84

## 4 结语

本文使用了卷积神经网络模型来提取图像的特征,利用流形排序算法对图像数据进行检索,返回相似图像的排序结果。实验结果表明深度图像特征和流形排序对于图像检索行之有效,深度学习提取的图像特征对于现有的相似性度量算法有着明显提升,表明深度学习在图像检索等方面有着巨大的潜力,深度图像特征能够提升检索效果。通过流形排序和余弦距离两者之间的对比,流形排序算法在深度图像特征上有着较好的效果。当然本文方法也存在一些不足之处:本文使用的卷积神经网络模型是已经训练好的,针对不同的图像数据集提取的图像特征不一定都有较好的检索效果。下一步的工作将会从重新训练或者微调卷积神经网络模型,对提取的图像特征进行降维处理以及用一种更好的相似性度量方法等方面展开,以进一步提高图像检索的准确率。

### 参考文献:

- [1] WU L, HOI S C H, YU N. Semantics-preserving bag-of-words models and applications [J]. IEEE Transactions on Image Processing, 2010, 19(7): 1908–1920.
- [2] YANG J, JIANG Y G, HAUPTMANN A G, et al. Evaluating bag-of-visual-words representations in scene classification [C]// Proceedings of the 2007 International Workshop on Workshop on Multimedia Information Retrieval. New York: ACM, 2007: 197–206.
- [3] LOWE D G. Object recognition from local scale-invariant features [C]// Proceedings of the 1999 IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 1999: 1150–1157.
- [4] BAY H, ESS A, TUYTELAARS T, et al. Speeded-Up Robust Features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346–359.
- [5] SCHMIDHUBER J. Deep learning in neural networks: an overview [J]. Neural Networks, 2015, 61: 85–117.
- [6] WAN J, WANG D, HOI S C H, et al. Deep learning for content-based image retrieval: a comprehensive study [C]// Proceedings of the 2014 ACM International Conference on Multimedia. New York: ACM, 2014: 157–166.
- [7] WU P, HOI S C H, XIA H, et al. Online multimodal deep similarity learning with application to image retrieval [C]// Proceedings of the 21st ACM International Conference on Multimedia. New York: ACM, 2013: 153–162.

## 参考文献:

- [1] SCHROEDER W, ZARGE J, LORENSEN W. Decimation of triangle meshes [J]. *Computer Graphics*, 1992, 26(2): 65–70.
- [2] 周元峰, 张彩明, 贺平. 体积平方度量下的特征保持网格简化方法[J]. *计算机学报*, 2009, 32(2): 203–212. (ZHOU Y F, ZHANG C M, HE P. Feature preserving mesh simplification algorithm based on square volume measure [J]. *Chinese Journal of Computers*, 2009, 32(2): 203–212.)
- [3] 张霞, 段黎明, 刘璐. 保持特征的高质量三角网格简化方法[J]. *计算机集成制造系统*, 2014, 3(3): 486–493. (ZHANG X, DUAN L M, LIU L. High quality triangular mesh simplification with feature preserving [J]. *Computer Integrated Manufacturing Systems*, 2014, 3(3): 486–493.)
- [4] 易兵, 刘振宇, 谭建荣. 边界特征保持的网格模型分级二次误差简化算法[J]. *计算机辅助设计与图形学学报*, 2012, 24(4): 428–434. (YI B, LIU Z Y, TAN J R. New quadric metric for simplifying meshes to retain the feature edge [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2012, 24(4): 428–434.)
- [5] 王海玲, 王建, 印桂生, 等. 多特征融合的网格模型简化方法[J]. *计算机应用*, 2013, 33(11): 3167–3171. (WANG H L, WANG J, YIN G S, et al. Multi-feature fusion method for mesh simplification [J]. *Journal of Computer Applications*, 2013, 33(11): 3167–3171.)
- [6] GARLAND M, HECKBERT P. Surface simplification using quadric error metrics [C]// *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. New York: ACM, 1997: 209–216.
- [7] LI L, HE M, WANG P. Mesh simplification algorithm based on absolute curvature-weighted quadric error metrics [C]// *Proceedings of 2010 the 5th IEEE Conference on Industrial Electronics and Applications*. Piscataway, NJ: IEEE, 2010: 399–403.
- [8] 赵晔, 周畅, 王昌. 一种随机采样的特征保持的网格简化算法[J]. *计算机科学*, 2011, 38(5): 249–251. (ZHAO Y, ZHOU C, WANG C. Feature preserved mesh simplification algorithm based on stochastic sampling [J]. *Computer Science*, 2011, 38(5): 249–251.)
- [9] UNGVICHIAN V, KANONGCHAIYOS P. Mesh simplification method using principal curvatures and directions [J]. *Computer Modeling in Engineering & Sciences*, 2011, 77(3): 201–219.
- [10] TSENG J, LIN Y. 3 D Surface simplification based on extended shape operator [J]. *WSEAS Transactions on Computers*, 2013, 12(8): 320–330.
- [11] LEE Y, MARSHALL D. Curvature based normalized 3D component facial image recognition using fuzzy integral [J]. *Applied Mathematics and Computation*, 2008, 205(2): 815–823.
- [12] 同济大学数学系. 高等数学[M]. 6版. 北京: 高等教育出版社, 2007: 170–174. (Department of Mathematics of Tongji University. *Higher mathematics* [M]. 6th ed. Beijing: Higher Education Press, 2007: 170–174.)
- [13] CIGNONI P, ROCCHINI C, SCOPIGNO R. Metro: measuring error on simplified surfaces [J]. *Computer Graphics Forum*, 1998, 17(2): 167–174.
- [14] CHEN C, CHENG K. A sharpness-dependent filter for recovering sharp features in repaired 3D mesh models [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2008, 14(1): 200–212.

## Background

This work is partially supported by the Nanning Scientific Research and Technological Development Project (20156197), the Scientific Research Project of Guangxi Higher Education Institutions of China (KY2015YB008).

**LIU Jun**, born in 1974, Ph. D., associate professor. His research interests include virtual reality, parallel computing.

**FAN Hao**, born in 1990, M. S. candidate. His research interests include graphics and image processing, distributed simulation.

**SUN Yu**, born in 1981, Ph. D., lecturer. Her research interests include intelligence computing.

**LU Xiangyan**, born in 1973, M. S., associate professor. Her research interests include parallel computing.

**LIU Yan**, born in 1972, Ph. D., associate research fellow. Her research interests include computer simulation.

(上接第534页)

- [8] KRIZHEYSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]// *Proceedings of the 26th Annual Conference on Neural Information Processing Systems*. Lake Tahoe, Nevada: [s. n.], 2012: 1097–1105.
- [9] VEDALDI A, LENC K. MatConvNet — convolutional neural networks for MATLAB [EB/OL]. [2015-06-21]. <http://arxiv.org/pdf/1412.4564.pdf>.
- [10] CHATFIELD K, SIMONYAN K, VEDALDI A, et al. Return of the devil in the details: delving deep into convolutional nets [EB/OL]. [2014-11-05]. <http://arxiv.org/pdf/1405.3531.pdf>.
- [11] BABENKO A, SLESAREV A, CHIGORIN A, et al. Neural codes for image retrieval [C]// *ECCV 2014: Proceedings of the 13th European Conference on Computer Vision*. Berlin: Springer, 2014: 584–599.
- [12] DONAHUE J, JIA Y, VINYALS O, et al. DeCAF: a deep convolutional activation feature for generic visual recognition [EB/OL]. [2015-05-10]. <http://arxiv.org/abs/1310.1531v1>.
- [13] ZHOU D, WESTON J, GRETTON A, et al. Ranking on data manifolds [C]// *NIPS 2003: Proceedings of the 2003 Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2004, 16: 169–176.
- [14] XU B, BU J, CHEN C, et al. Efficient manifold ranking for image retrieval [C]// *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York: ACM, 2011: 525–534.
- [15] OLIVA A, TORRALBA A. Modeling the shape of the scene: a holistic representation of the spatial envelope [J]. *International Journal of Computer Vision*, 2001, 42(3): 145–175.
- [16] CHECHIK G, SHARMA V, SHALIT U, et al. Large scale online learning of image similarity through ranking [J]. *Journal of Machine Learning Research*, 2009, 11(2): 1109–1135.

## Background

This work is partially supported by the National Natural Science Foundation of China (61003127, 61373109).

**LIU Bing**, born in 1992, M. S. candidate. His research interests include deep learning, image processing.

**ZHANG Hong**, born in 1979, Ph. D., professor. Her research interests include cross-media retrieval, machine learning, data mining.