·图形图像处理· 文章编号:1000-3428(2018)06-0200-07 文献标志码:A 中图分类号:TP391

# 嵌入标记信息的铁路扣件状态检测主题模型

欧 阳,罗建桥,李柏林,李 爽

(西南交通大学 机械工程学院,成都 610031)

**摘 要:** 针对潜在狄利克雷分布(LDA)模型忽略特征单词明确性的问题,提出一种嵌入标记信息的主题模型 WL_LDA。设计一种基于 SIFT 特征点约束单方向 LBP 图像的方法。运用该方法获取图像的纹理结构,对视觉单词进行标记。将标记信息嵌入到 LDA 中,利用单词和标记的二维直方图推导图像的主题分布。通过运用该主题分布训练分类器,完成铁路扣件的状态检测。实验结果表明,与 LDA 主题模型相比,各扣件在主题空间中的区分度增加4.5% ~15% ,与现有 PCA、DF 等方法相比,漏检率和误检率明显降低,具有较好的分类性能。

**关键词:** 图像语义分析;潜在狄利克雷分布;视觉单词;SIFT 特征;单词标记;主题模型

**中文引用格式:**欧 阳,罗建桥,李柏林,等. 嵌入标记信息的铁路扣件状态检测主题模型[J]. 计算机工程,2018,44(6):200-206.

**英文引用格式:**OU Yang,LUO Jianqiao,LI Bailin,et al. Topic model of railway fastener state detection embedded with tag information[J]. Computer Engineering,2018,44(6):200-206.

# Topic Model of Railway Fastener State Detection Embedded with Tag Information

OU Yang,LUO Jianqiao,LI Bailin,LI Shuang

(School of Mechanical Engineering,Southwest Jiaotong University,Chengdu 610031,China)

**【Abstract】** Aiming at the explicit problem of ignoring the characteristic words in the Latent Dirichlet Allocation (LDA) model,a topic model WL_LDA embedded tag information is proposed. A method for constraining unidirectional LBP images based on SIFT feature points is designed to obtain the texture structure of the image and the visual word is marked. By embedding the tag information in the LDA,the topic distribution of the image is derived using the two-dimensional histograms of words and tags. This topic distribution is used to train the classifier to complete the state detection of railway fasteners. Experimental results show that,compared with the LDA topic model,the differentiation degree of various types of fasteners in the topic space increases by 4.5% to 15% . Compared with the existing methods such as PCA and DF,the rate of missing detection and false detection is significantly reduced,it has better classification performance.

**【Key words】** image semantic analysis;Latent Dirichlet Allocation (LDA);visual word;SIFT feature;word marking;topic model

**DOI:**10.19678/j.issn.1000-3428.0047330

## 0 概述

铁路扣件是组成铁轨线路的必要部件,为保障线路的安全,对扣件状态的检测显得至关重要。在国外,对扣件状态的检测已有了相关研究[1]。近年来,国内的研究人员针对我国扣件的类型,提出了一些检测算法。文献[2]提出一种钢轨扣件螺母缺失检测系统,该系统通过主成分分析(Principal Component Analysis,PCA),根据全局特征向量的距离对扣件进行分类,忽略了局部特征,检测效果欠佳。文献[3]提出一种用局部特征检测扣件状态的方法,虽改善了对失效扣件的检测,但却误检大量正常扣件。文献[4]采用模板匹配对扣件进行分类,由于在实际工况下的扣件形状各异,因此匹配度不理想,误检率很高。文献[5]提出一种基于扣件局部特征和语义信息的扣件检测模型,改善了扣件的描述能力。

**基金项目:**四川省科技支撑计划项目(2016GZ0194)。

**作者简介:**欧 阳(1991—),男,博士研究生,主研方向为图像处理、模式识别;罗建桥,博士研究生;李柏林,教授、博士;李 爽,硕士研究生。

**收稿日期:**2017-05-24 **修回日期:**2017-07-04 **E-mail:**you91swjtu@163.com

底层特征在描述扣件轮廓方面存在不足,用其表达图像时,会造成较大的语义鸿沟。文献[6]提出对图像底层特征进行语义分析极其重要。作为一种具有代表性的主题模型,潜在狄利克雷分布(Latent Dirichlet Allocation,LDA)[7]在文本分类、图像标注等领域得到了广泛的应用。如文献[8]采用附带类别标签的 LDA 模型用于文本分类。文献[9]针对传统基于多项式分布的主题模型不能较好地刻画文档中词汇突发的现象,提出了一种面向词汇突发的 Dirichlet 组合多项式连续时间主题模型。文献[10]通过融合语义主题进行图像自动标注。文献[11]指出,纹理结构是人类认知目标的关键,如若忽略将造成扣件图像内容的描述能力不足。文献[12]在进行扣件分类时,融入了图像的纹理结构信息,取得了较好的结果。

在上述研究的基础上,本文提出一种嵌入单词标记信息的 LDA(Word Labeled LDA,WL_LDA)检测模型。该模型根据图像在高斯空间计算的尺度不变局部特征点(SIFT)[13],约束单方向 LBP 图像,获取纹理结构图像,对视觉单词进行标记,图像由单词-标记二维直方图表示,并将标记信息嵌入到 LDA 生成模型中,图像的主题由单词-标记二维直方图推导。

## 1　单词-标记二维直方图的获取

铁路扣件图像的组成包括扣件结构和背景。本文设计一种用 SIFT 特征点约束单方向 LBP 图像获取扣件图像纹理结构的方法。根据视觉单词在纹理结构图中对应位置的像素强度关系,对视觉单词进行标记。通过标记,解决视觉单词的歧义性问题。

### 1.1　SIFT 特征点的提取

在图像的线性尺度空间中,寻找局部极值点作为扣件结构分布的重要信息。文献[13]提出的 SIFT 特征,具有对旋转、尺度缩放、亮度变化保持不变的特点,鲁棒性强,对噪声也有一定的稳定性。本文采用与文献[13]相同的参数设置,即 $r=8$。图 1 所示的白色点是 SIFT 特征点的示例。

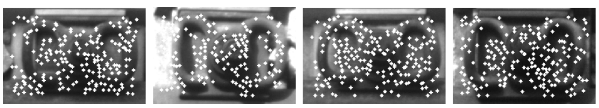


**图 1　扣件图像的 SIFT 特征点示例**

从图 1 可以看出,在不同光照、不同视角等条件下,扣件弹条周围均分布有 SIFT 点。因此,SIFT 特征点的分布能较好地反映扣件结构的分布。

### 1.2　单方向 LBP 图像扣件纹理结构

受文献[14]启发,在 8 个方向上分别提取扣件图像的 LBP 特征。定义每个像素点在某方向上的 LBP 编码为:

$$LBP_c^p(x,y)=\sigma(g_p(x,y)-g_c(x,y)) \tag{1}$$

其中,$g_c(x,y)$是中心像素的灰度值,$g_p(x,y)$是中心像素 $g_c$ 在 $P$ 方向上邻域点的灰度值,$LBP_c^p(x,y)$是中心像素 $g_c$ 在 $P$ 方向上的纹理编码值,$\sigma(t)$是状态判别函数,当 $t\geqslant 0$ 时,值为 1,否则为 0。

$g_p(x,y)$的定义如下:

$$g_p(x,y)=\begin{cases} g_c(x,y+R), p=0^\circ \\ g_c(x-R,y+R), p=45^\circ \\ g_c(x-R,y), p=90^\circ \\ g_c(x-R,y-R), p=135^\circ \end{cases} \tag{2}$$

其中,$R$ 为中心像素 $g_c$ 的邻域半径。

一幅扣件图像及其 4 个不同方向上的单方向 LBP 图像如图 2(a)~图 2(e)所示,其中,$R=4$ 为优化值,不在此赘述。从图 2(a)~图 2(e)可以看出,90°方向的 $LBP_c^{90^\circ}$ 图(图 2(d))能较好地描绘扣件的结构,以此为基础进行形态学开操作,得到扣件的纹理结构 $I_{\mathrm{pt}}$,如图 2(f)所示。

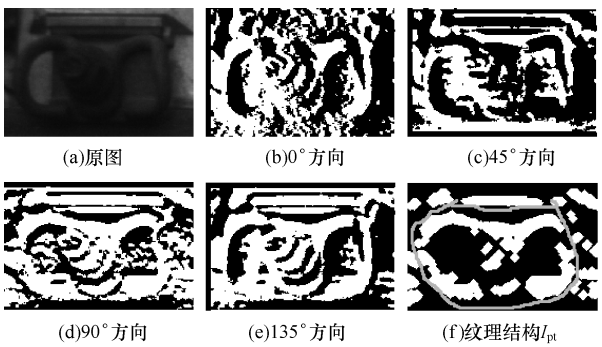



**图 2　单方向 LBP 扣件纹理结构**

### 1.3　单方向 LBP 纹理结构的约束

从图 2(f)中可以看出,纹理结构图 $I_{\mathrm{pt}}$ 虽然较为完整地保留了扣件的结构要素,但存在较大的背景干扰信息,如图 2(f)中圆圈区域外所示。若直接用其进行单词的标记,将会使标记的单词产生歧义。由 1.1 节可知,扣件图像 SIFT 特征点的分布能较好地反映扣件结构的分布。利用这一关系,对扣件纹理结构图像 $I_{\mathrm{pt}}$ 进行约束,得到最终的纹理结构图像 $I_{\mathrm{PT}}$,如图 3 所示。

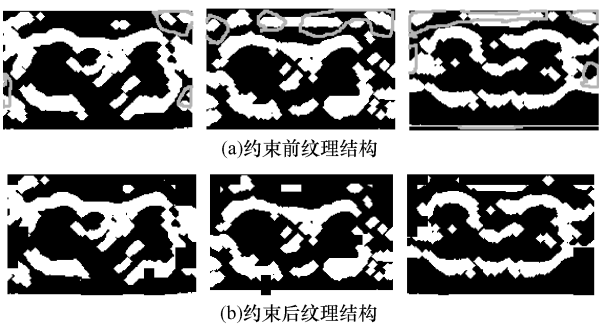



**图 3　约束前后纹理结构的对比**

从图 3 可以看出,约束后的纹理结构图像 $I_{FT}$ 与约束前的纹理结构图像 $I_{pt}$ 相比,去掉了较多的背景干扰信息,如图 3(a)中圆圈区域内所示。

获取纹理结构图像 $I_{FT}$ 的过程如下:采用均匀网格采样方法,将纹理结构图像 $I_{pt}$ 划分为 $N$ 个子块,子块大小为 20×20,依次按“从左到右”“从上到下”获取图像子块的像素强度信息;同样地,将原图划分为 $N$ 个子块,子块大小为 20×20,依次按“从左到右”“从上到下”获取图像子块的 SIFT 特征点信息。当对应子块的像素强度满足结构条件,且子块内 SIFT 特征点满足条件时,如式(3)所示,则认为其为目标,将其保留;否则,将视为背景移除。

$$\sum_x \sum_y I(x,y) > 0 \&\& num_SIFT \geqslant 2 \tag{3}$$

其中,$\sum_x \sum_y I(x,y)$ 表示纹理结构图像 $I_{pt}$ 在子块里的像素强度(灰度)值的和,$num_SIFT$ 表示在原图对应子块的 SIFT 特征点个数。

### 1.4 纹理结构图像 $I_{FT}$ 标记的视觉单词

获取图像视觉单词直方图表示的一般过程为:提取特征;通过聚类法则,形成词典;根据词典编码特征。图像的传统视觉单词表示是一个一维的直方图,如图 4 所示(以 LBP 作为底层特征)。

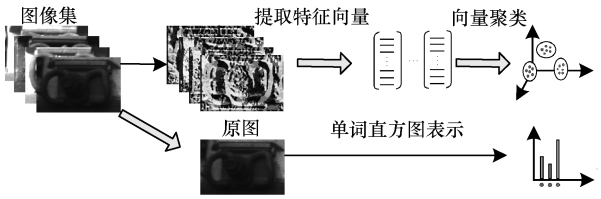
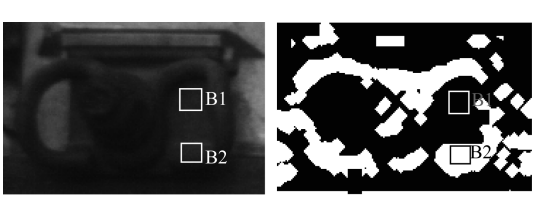



**图 4 传统视觉单词直方图描述图像的过程**

传统视觉单词的问题在于:根据聚类法则,若不同图像区域的特征相似,则区域的单词编码相同,因而造成单词的歧义性。如图 5(a)所示,在原始图像中,由于背景区域 B1、扣件弹条区域 B2 的特征相似,则 2 个区域的单词编码相同,$w_{上} = w_{下}$,采用 $w$ 无法区分 B1 和 B2。

为了解决歧义性问题,在结构图像中对单词进行标记。如图 5(b)所示,在结构图像中,B1 的标记为 $s=0$,代表背景;B2 的标记为 $s=1$,代表扣件结构。通过标记,即可区分 B1 和 B2。




(a)原始图像编码单词 (b)结构图像标记单词

**图 5 单词歧义性的克服**

获取一幅扣件图像视觉单词标记如算法 1 所示。

**算法 1** 获取单词标记直方图

**输入** 训练图像

**输出** 单词-标记直方图

**步骤 1** 搭建视觉词典。首先,获取训练图像中所有子块的局部特征向量;然后,进行 K-均值聚类,得到视觉词典 $V=\{v_1,v_2,\cdots,v_v\}$,$v_i$ 表示视觉词典中第 $i$ 个视觉单词,词典的单词容量为 $V$。

**步骤 2** 初始化一个带标记的单词直方图。

$$H(v_d,s)=0,d=1,2,\cdots,v,s=0,1 \tag{4}$$

**步骤 3** 获取子块特征 $n$ 的单词编号 $v_i$。

$$v_i=\arg\min_i \{dis(n-v_i)\},i=1,2,\cdots,v \tag{5}$$

**步骤 4** 对单词进行标记。对于原图中子块 $n$ 的位置,定位在纹理结构 $I_{FT}$ 中对应的子块 $n^{FT}$,按 $n^{FT}$ 在对应子块中的像素强度占比,如式(6)所示。

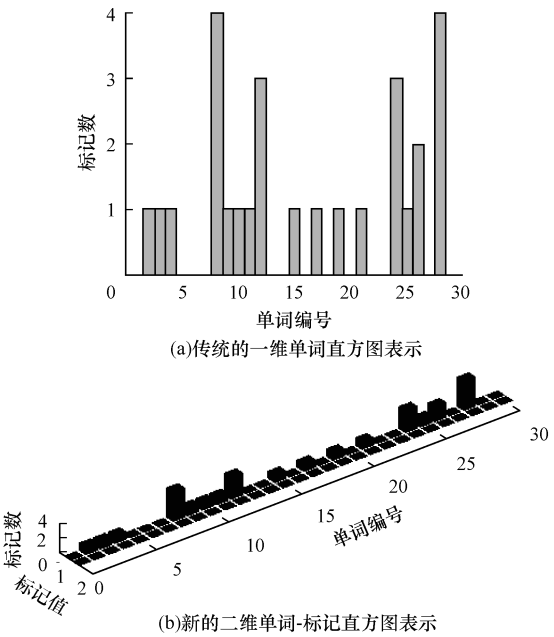
$$s=\begin{cases}1, \dfrac{\text{sum}\{n^{FT}(x,y)=1\}}{\text{sum}\{n^{FT}(x,y)\}} \geqslant k \\ 0, \text{其他}\end{cases} \tag{6}$$

其中,$n^{FT}(x,y)$ 表示在纹理结构图像 $I_{FT}$ 中位于 $(x,y)$ 处像素点的灰度值,其取值为 0 或 1,sum 为求和符号,$k$ 为优化常数,本文取 $k=0.8$。

**步骤 5** 更新直方图:$H(v_i,s)=H(v_i,s)+1$。

**步骤 6** 循环,直至对一幅图像的所有子块标记结束。

根据以上分析,标记信息克服了单词的歧义性问题,突出了单词的语义,用标记的单词来描述图像显然比未标记的单词更加清晰。图 6 为传统的一维单词直方图表示和新的二维单词-标记直方图表示,由于空间限制,仅列出了前 30 号单词的分布。




(a)传统的一维单词直方图表示

(b)新的二维单词-标记直方图表示

**图 6 2 种直方图表示方法**

## 2　嵌入标记信息的 WL_LDA 模型

本文提出的 WL_LDA 模型研究单词-标记直方图背后的主题语义,用嵌入标记信息的生成模型推导图像主题,并用其描述扣件。本文方法是以 LDA 模型为基础并引入标记变量 $s$,单词 $w$ 的生成由主题 $z$ 和标记 $s$ 共同决定。

### 2.1　概率图模型及参数

受经典 LDA 主题模型[7]的启发,沿用其部分参数,并引入新的变量:标记变量 $s$。按照 LDA 主题模型的策略定义变量,下面仅列出与 LDA 模型不同的变量。

$s_i$:单词的标记变量,其取值为 $t$,$t$ 取 0 或 1。

$\boldsymbol{\psi}_{z_i}$:标记分布参数。

$\boldsymbol{\gamma}$:标记分布参数 $\boldsymbol{\psi}_{z_i}$ 的先验。

$\boldsymbol{\varphi}_{z_i,s_i}$:单词分布参数;单词的分布参数由主题 $z$ 与标记 $s$ 共同决定。

根据定义的参数,得出本文提出的 WL_LDA 概率图模型,如图 7 所示。

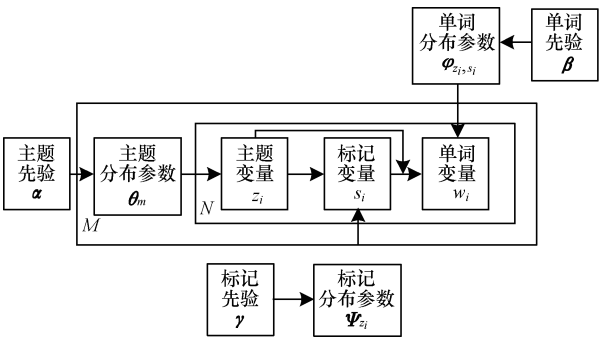



**图 7　WL_LDA 概率图模型**

### 2.2　WL_LDA 模型的物理过程及参数估计

将图 7 中 WL_LDA 的概率图模型分解为 3 个主要的物理过程:由主题先验到主题变量;由标记先验到标记变量;由单词先验到单词变量。与 LDA 不同的物理过程如下所示。

由标记先验到标记变量 $\boldsymbol{\gamma}\rightarrow\boldsymbol{\psi}_k\rightarrow s_i$:

$$\boldsymbol{\psi}_k \sim \mathrm{Dir}(\boldsymbol{\psi}_k \mid \boldsymbol{\gamma}) \tag{7}$$

$$s_i \sim \mathrm{mult}(s_i \mid \boldsymbol{\psi}_k) \tag{8}$$

由单词先验到单词变量 $\boldsymbol{\beta}\rightarrow\boldsymbol{\varphi}_{kt}\rightarrow w_i$:

$$\boldsymbol{\varphi}_{kt} \sim \mathrm{Dir}(\boldsymbol{\varphi}_{kt} \mid \boldsymbol{\beta}) \tag{9}$$

$$w_i \sim \mathrm{mult}(w_i \mid \boldsymbol{\varphi}_{kt}) \tag{10}$$

其中,Dir 表示狄利克雷分布,mult 表示多项分布,$s_i$ 和 $w_i$ 是可见变量。

在 WL_LDA 模型中,有 3 个随机变量:主题 $z$,标记 $s$,单词 $w$。

那么,图像集可以描述为 $\{\boldsymbol{z},\boldsymbol{s},\boldsymbol{w}\}$,整个图像集的似然函数为:

$$likelihood = p(\boldsymbol{z},\boldsymbol{s},\boldsymbol{w} \mid \boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta}) \tag{11}$$

根据 3 个物理过程及概率图可知:当主题变量确定时,标记变量和主题先验的关系为条件独立,单词变量与主题先验的关系也是条件独立;当标记变量确定时,单词变量与标记先验的关系为条件独立。因此,式(11)可以分解为式(12)。

$$\begin{aligned} likelihood = p(\boldsymbol{z},\boldsymbol{s},\boldsymbol{w} \mid \boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta}) &= p(\boldsymbol{z} \mid \boldsymbol{\alpha}) \times \\ p(\boldsymbol{s} \mid \boldsymbol{z},\boldsymbol{\gamma}) &\times p(\boldsymbol{w} \mid \boldsymbol{z},\boldsymbol{s},\boldsymbol{\beta}) \end{aligned} \tag{12}$$

由此得到主题、标记、单词的联合分布如式(13)所示。

$$\begin{aligned} p(\boldsymbol{z},\boldsymbol{s},\boldsymbol{w} \mid \boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta}) &= \prod_{m=1}^{M} \frac{\Delta(\boldsymbol{n}_m + \boldsymbol{\alpha})}{\Delta(\boldsymbol{\alpha})} \cdot \\ \prod_{k=1}^{K} \frac{\Delta(\boldsymbol{n}_k + \boldsymbol{\gamma})}{\Delta(\boldsymbol{\gamma})} &\cdot \prod_{k=1}^{K} \prod_{t=0}^{1} \frac{\Delta(\boldsymbol{n}_{k,t} + \boldsymbol{\beta})}{\Delta(\boldsymbol{\beta})} \end{aligned} \tag{13}$$

其中,$\boldsymbol{n}_m$ 表示在第 $m$ 张图像中每个主题出现的次数,$\boldsymbol{n}_m = \{n_m^k\}$,$k = 1 \sim K$,$\boldsymbol{n}_k$ 表示主题 $k$ 下每一类标记出现的次数,$\boldsymbol{n}_k = \{n_k^t\}$,$t = 0 \sim 1$,$\boldsymbol{n}_{k,t}$ 表示当主题变量 $z = k$、标记变量 $s = t$ 时,每种单词出现的次数,$\boldsymbol{n}_{k,t} = \{n_{k,t}^v\}$,$v = 1 \sim V$。

受文献[15]的启发,利用吉布斯采样原理,可得主题 $z$ 的采样公式为:

$$\begin{aligned} p(z_i = k \mid \boldsymbol{z}_{\neg i},\boldsymbol{s},\boldsymbol{w},\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta}) &\propto \frac{n_{m,\neg i}^k + \alpha_k}{\sum_{z=1}^{K}(n_{m,\neg i}^z + \alpha_k)} \cdot \\ \frac{n_{k,\neg i}^t + \gamma_t}{\sum_{t=0}^{1}(n_{k,\neg i}^t + \gamma_t)} &\cdot \frac{n_{k,t,\neg i}^v + \beta_v}{\sum_{v=1}^{V}(n_{k,t,\neg i}^v + \beta_v)} \end{aligned} \tag{14}$$

其中,$\alpha_k$ 表示主题先验 $\boldsymbol{\alpha}$ 的第 $k$ 个主题出现的先验概率,同理可得 $\gamma_t$ 与 $\beta_v$ 的含义。

由于省略了常数项 $p(s_i = t, w_i = v)$,因此引入正比符号。

$\neg i$ 表示不统计当前采样数据,如 $n_{m,\neg i}^k$ 表示第 $m$ 张图像中主题 $k$ 出现的次数减 1。

从式(14)可以看出,对图像主题的推导,包含了标记信息 $s$,即图像的主题是由 1.4 节中提出的单词-标注二维单词直方图推导,不再由单词 $w$ 唯一决定。

综上所述,本文提出的检测框图如图 8 所示。

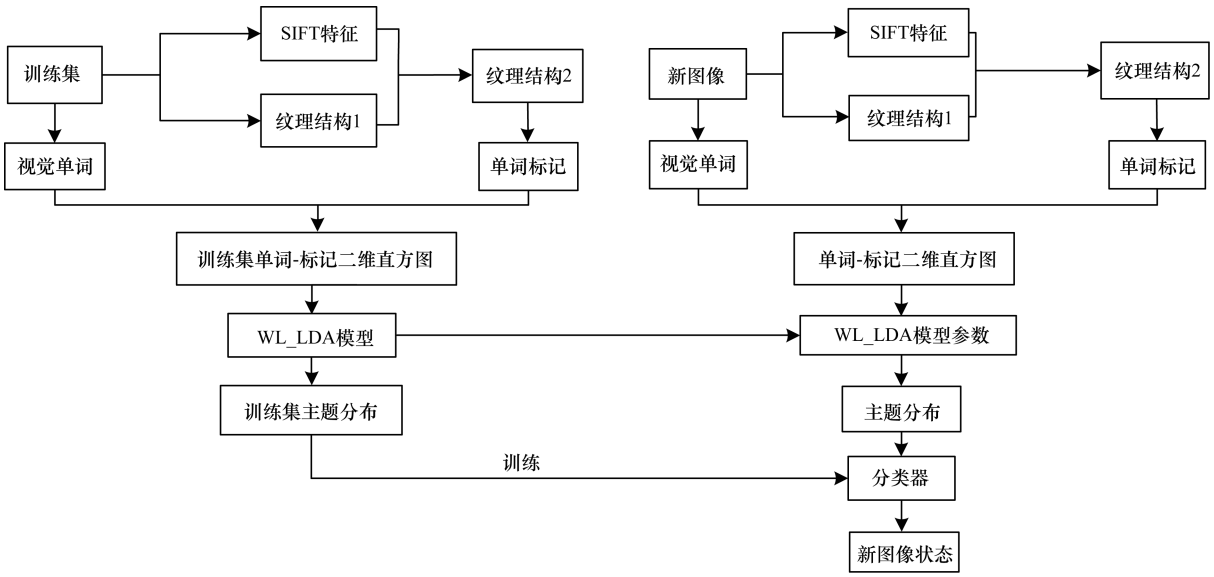
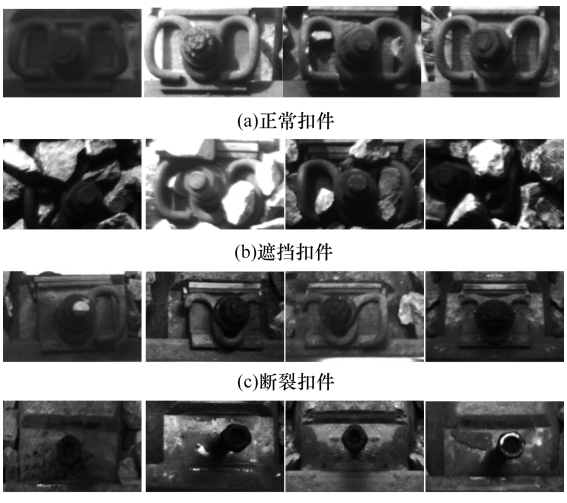



**图 8 本文检测算法框图**

## 3 实验结果与分析

### 3.1 实验环境

为测试本文方法的检测性能,建立一个扣件图像样本集,扣件类别包括正常、遮挡、断裂以及丢失,除正常扣件外,其余的统称为失效扣件。在实际工况下,铁路沿线的扣件绝大部分都是正常的,失效的扣件比例极小。为正确评价各扣件检测方法的性能,从大量扣件图像中选择正常扣件 700 张、遮挡扣件 400 张、断裂扣件 200 张、丢失扣件 100 张作为样本库。图像均为 120 × 180 像素的灰度图。部分实验图像样本如图 9 所示。



(a)正常扣件

(b)遮挡扣件

(c)断裂扣件

(d)丢失扣件

**图 9 部分实验图像样本**

从样本集中随机抽取 100 张正常扣件、100 张遮挡扣件、100 张断裂扣件以及 50 张丢失扣件作为训练集,其余图像作为验证集。实验 PC 处理器为 Inter(R) Core(TM) i5-6500 CPU @ 3.20 GHz,内存 4.00 GB,在 Matlab2014a 环境下进行实验。为减少实验的随机性,采用交叉验证的方式,从样本集中重复随机抽取 10 次,即得到 10 组训练集和验证集。取 10 组数据的平均值作为本文实验结果。其中,扣件漏检是指失效扣件未被检查出来,视为正常扣件,扣件误检是指有的正常扣件被视为失效扣件检测出来,漏检率 = 漏检量/失效扣件总量 × 100%,误检率 = 误检量/正常扣件总量 × 100%。

### 3.2 结果分析

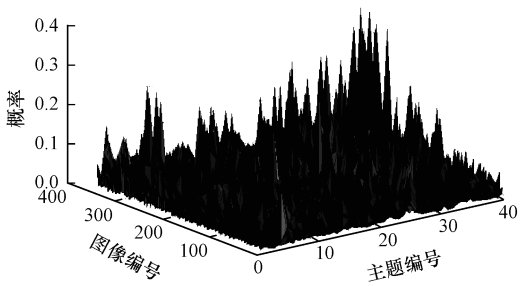
主题模型与扣件状态检测实验方法如下:

1)WL_LDA 主题模型能力评估实验

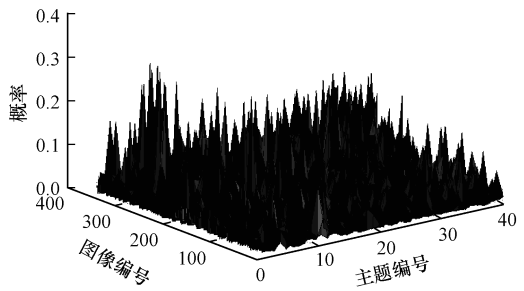
测试 WL_LDA 模型所推导主题的区分性能与 LDA 模型所推导的主题进行对比。首先,分别采用两种模型训练图像的主题;然后,比较各类扣件主题分布的距离,并以此衡量模型的区分性能,值越大,区分性能越好。将图像划分成大小为 20 × 20 的子块,移动步长为 10,在子块上提取 LBP 特征作为底层特征。

实验参数的设置分别为:视觉词典的容量采用 $V = 200$;主题的容量 $K = 40$;标记数 $t = 2$;先验参数 $\alpha = 50/K, \gamma = 0.1, \beta = 0.1$。优化参数不在此赘述。选择支持向量机库(Library for Support Vector Machine,LIBSVM)作为分类器。

推导出的主题分布如图 10 所示。其中,扣件图像的状态为:1 ~ 100 正常,101 ~ 200 遮挡,201 ~ 300 断裂,301 ~ 350 丢失。




(a)LDA模型下的主题分布




(b)WL_LDA模型下的主题分布

**图 10　扣件状态的主题分布**

通过主题模型推导出主题分布之后,每一张扣件图像都由一个主题分布矩阵来表达。为了测量这些以主题分布直方图表达的图像之间的差异性,采用式(15)计算不同类别扣件主题之间的距离。

$$Dis(H,M)=\sum_{i=1}^{L}(\bar{H}_i-\bar{M}_i)^2 \tag{15}$$

其中,$H$、$M$ 分别表示扣件的 2 个不同类别,$L$ 表示主题直方图的维度,$\bar{H}$ 表示该类扣件下,其所有主题分布直方图的均值直方图。

表 1 给出了在 WL_LDA 模型和 LDA 模型下不同类别扣件的距离及增长率。

**表 1　2 种模型主题空间下的距离对比**

| 扣件状态 | 主题空间距离 | | 增长率/% |
|---|---|---|---|
| | LDA 模型 | WL_LDA 模型 | |
| 正常-遮挡 | 0.088 | 0.092 | 4.5 |
| 正常-断裂 | 0.120 | 0.138 | 15.0 |
| 正常-丢失 | 0.116 | 0.126 | 8.6 |

从表 1 可以看出,WL_LDA 模型推导出的主题相较于 LDA 模型,其类间距离均有较大的提高,说明了该模型在进行扣件状态分类时,其区分能力更强。特别地,由于在 LDA 模型中嵌入了标记信息,解决了单词的歧义性问题,能够区分扣件弹条与背景,使正常扣件与断裂扣件的区分度增大,这与实验结果相吻合。

2)WL_LDA 主题模型扣件状态检测实验

实验 1 充分说明了 WL_LDA 模型推导主题的区分能力。为准确评价本文方法的扣件状态检测性能,与目前几种语义方法进行对比实验,即经典的词包模型(BOW)、LDA 模型以及空间金字塔 LDA 模型(S_LDA)[16]。BOW 模型是以视觉单词表示图像;LDA 模型是以视觉单词为基础,推导主题来表达图像;S_LDA 模型是在 LDA 模型的基础上,加入了空间结构要素。

上述几种方法分别以 LBP 和 HOG 为底层特征,其中 LBP 的具体取值与实验 1 相同,HOG 的 cell 尺寸为 5×5,梯度方向为 9 个。S_LDA 模型的金字塔层数取 2,主题空间中的特征长度取 40×(1+4)=200。分类器与实验 1 相同。实验结果如表 2 所示。

**表 2　不同语义的实验结果**

| 实验 | 语义 | 漏检率/% | 误检率/% | 时间消耗/(张·$s^{-1}$) |
|---|---|---|---|---|
| 1 | HOG + BOW | 4.1 | 8.4 | 0.68 |
| 2 | HOG + LDA | 3.5 | 6.6 | 0.82 |
| 3 | HOG + S_LDA | 1.7 | 9.7 | 1.02 |
| 4 | HOG + WL_LDA | 1.8 | 6.3 | 1.15 |
| 5 | LBP + BOW | 3.2 | 7.5 | 0.13 |
| 6 | LBP + LDA | 2.7 | 6.1 | 0.21 |
| 7 | LBP + S_LDA | 1.8 | 7.8 | 0.24 |
| 8 | LBP + WL_LDA | 1.5 | 4.3 | 0.30 |

实验 1 ~ 实验 4 结果表明,用 HOG 特征作为底层特征,本文提出的 WL_LDA 模型,在漏检率方面,与最低的 S_LDA 模型相近;在误检率方面,低于其他所有的语义模型。原因是:对于 BOW 模型,将 HOG 特征转换为视觉单词,是一种底层语义到中层语义的总结;对于 LDA 模型,在中层语义中推导图像主题,是在 BOW 模型的基础上进一步总结语义信息,降低误检率和漏检率;对于 S_LDA 模型,虽然通过加入空间结构要素降低漏检率,但是由于正常扣件并非总是位于图像正中,因此正常扣件的主题分布差别较大,误检率较高。本文提出的 WL_LDA 模型,根据扣件图像的纹理结构对视觉单词进行标记,与扣件是否位于图像正中无关,整合得到的单词-标记联合分布克服了单词的歧义性问题,推导的主题更加准确,能够降低漏检并减少误检。实验 5 ~ 实验 8的结果表明,用 LBP 特征作为底层特征,结果与实验 1 ~ 实验 4 类似,充分说明了本文提出的 WL_LDA 模型,其较优秀的性能与底层特征的选取无关。

综合以上 8 组实验可以看出,本文算法在性能上较经典的 LDA 模型有较大的提高,对于铁路扣件状态的检测效果更佳。

3)与其他的扣件检测算法对比

为综合评价本文模型对铁路扣件的分类性能,与其他主要的扣件检测算法进行对比,包括文献[2]

的 PCA 算法、文献[3]的 DF 算法以及文献[4]的 LBP 和 HOG 算法。除 HOG 的参数设置与本文实验 2 相同外,其余实验设置与文献[12]的设置相同。分类器与实验 1 相同,各算法分类性能结果对比如表 3 所示。

**表 3 各扣件检测算法分类性能对比** %

| 实验 | 算法 | 漏检率 | 误检率 |
|---|---|---|---|
| 1 | PCA | 9.4 | 24.3 |
| 2 | DF | 4.9 | 14.6 |
| 3 | HOG | 5.2 | 9.2 |
| 4 | LBP | 3.6 | 9.6 |
| 5 | HOG + WL_LDA | 1.8 | 6.3 |
| 6 | LBP + WL_LDA | 1.5 | 4.3 |

通过对比表 3 中的 6 组实验可以看出,无论是特征维度较高的 LBP 方法和 HOG 方法,还是特征维度较低的 DF 方法和 PCA 方法,其漏检率和误检率都要高于本文方法。这是由于这些方法缺少语义分析,所获取的图像表示与真实图像存在较大的语义鸿沟,因此效果欠佳。

## 4 结束语

本文设计一种嵌入标记信息的铁路扣件状态检测主题模型——WL_LDA 模型,该模型使用视觉单词和标记信息来表示图像,使图像的内容更加丰富。用于铁路扣件状态检测,相较于现有的 LBP、HOG 等检测方法,其检测效果得到进一步提升,说明了本文方法的有效性。由于本文嵌入了标记信息,虽表达的图像内容更为清晰,但在构造分类器时,花费的时间更多,下一步将研究提出新的图像表示方法,以减少时间花费,提高算法实时性。

### 参考文献


[1] HSIEH H Y,CHEN N,LIAO C L. Visual recognition system of elastic rail clips for mass rapid transit systems[C]// Proceedings of ASME/IEEE 2007 Joint Rail Conference and Internal Combustion Engine Division Spring Technical Conference. Washington D. C.,USA:IEEE Press,2007:319-325.

[2] 王 凌,张 冰,陈锡爱. 基于计算机视觉的钢轨扣件螺母缺失检测系统[J]. 计算机工程与设计,2011,32(12):4147-4150.

[3] DOU Y,HUANG Y,LI Q,et al. A fast template matching-based algorithm for railway bolts detection[J]. International Journal of Machine Learning and Cybernetics,2014,5(6):835-844.

[4] YANG J,TAO W,LIU M,et al. An efficient direction field-based method for the detection of fasteners on high-speed railways[J]. Sensors,2011,11(8):7364-7381.

[5] 罗建桥,刘甲甲,李柏林,等. 基于局部特征和语义信息的扣件图像检测[J]. 计算机应用研究,2016,33(8):2514-2518.

[6] TANG J,ZHA Z J,TAO D,et al. Semantic-gap-oriented active learning for multilabel image annotation[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society,2012,21(4):2354-2360.

[7] BLEI D M,NG A Y,JORDAN M I. Latent dirichlet allocation[J]. Journal of Machine Learning Research,2003(3):993-1022.

[8] 李文波,孙 乐,黄瑞红,等. 基于 Labeled-LDA 模型的文本分类新算法[J]. 计算机学报,2008,31(4):620-627.

[9] 刘良选,黄梦醒. 一种面向词汇突发的连续时间主题模型[J]. 计算机工程,2016,42(11):195-201.

[10] 李志欣,施智平,李志清,等. 融合语义主题的图像自动标注[J]. 软件学报,2011,22(4):801-812.

[11] 刘甲甲,李柏林,罗建桥,等. 融合 PHOG 和 MSLBP 特征的铁路扣件检测算法[J]. 西南交通大学学报,2015,50(2):256-263.

[12] 罗建桥,刘甲甲,李柏林,等. 融合纹理结构的潜在狄利克雷分布铁路扣件检测模型[J]. 计算机应用,2016,36(2):574-579.

[13] LOWE D G. Distinctive image features from scale-invariant keypoints[M]. Boston,USA:Kluwer Academic Publishers,2004.

[14] 刘帅师,田彦涛,万 川. 基于 Gabor 多方向特征融合与分块直方图的人脸表情识别方法[J]. 自动化学报,2011,37(12):1455-1463.

[15] HEINRICH G. Parameter estimation for text analysis[EB/OL]. [2017-05-21]. http://m. 9512. net/read/672d64bd2900a276d22be3eb. html.

[16] WANG Y,LI J,GAO X. Latent feature mining of spatial and marginal characteristics for mammographic mass classification[J]. Neurocomputing,2014,144(1):107-118.


编辑 索书志