



(12) 发明专利申请

(10) 申请公布号 CN 105335758 A

(43) 申请公布日 2016. 02. 17

(21) 申请号 201510738843. 6

(22) 申请日 2015. 11. 03

(71) 申请人 电子科技大学

地址 611731 四川省成都市高新区(西区)西源大道 2006 号

(72) 发明人 李鸿升 胡欢 刘海军 曹滨 周辉

(74) 专利代理机构 成都弘毅天承知识产权代理有限公司 51230

代理人 徐金琼

(51) Int. Cl.

G06K 9/62(2006. 01)

G06K 9/46(2006. 01)

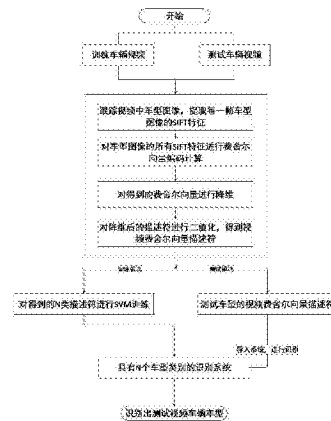
权利要求书2页 说明书6页 附图1页

(54) 发明名称

一种基于视频费舍尔向量描述符的车型识别方法

(57) 摘要

一种基于视频费舍尔向量描述符的车型识别方法,首先对用于训练的车辆视频,跟踪视频中每一种车型的图像,提取该车型每一帧图像的 SIFT 特征;然后对车型图像的所有 SIFT 特征进行费舍尔向量编码计算;接着对得到的费舍尔向量描述符进行 PCA 降维;再对降维后的描述符进行二值化,得到该车型的视频费舍尔向量描述符;将得到的所有描述符进行 SVM 训练,得到一个具有 N 个车型类别的识别系统;对用于测试的车辆视频,提取其视频费舍尔向量描述符,将其导入训练好的 SVM 分类器进行测试,识别出测试车辆视频的车型。



1. 一种基于视频费舍尔向量描述符的车型识别方法,包括以下步骤:

步骤 1:对用于训练的车辆视频,跟踪视频中每一种车型的图像,提取该车型每一帧图像的 SIFT 特征;

步骤 2:对提取的车型图像的所有 SIFT 特征进行费舍尔向量编码计算,得到费舍尔向量;

步骤 3:对得到的费舍尔向量描述符进行 PCA 降维;

步骤 4:对降维后的得到的费舍尔向量描述符进行二值化处理,得到该车型的视频费舍尔向量描述符;

步骤 5:将得到的所有视频费舍尔向量描述符进行 SVM 训练,得到一个具有 N 个车型类别的识别系统;

步骤 6:对测试车辆视频,同样提取视频中车辆图像的视频费舍尔向量描述符,将其导入步骤 5 中训练好的车型识别系统进行测试,识别出测试车辆视频的车型。

2. 根据权利要求 1 所述基于视频费舍尔向量描述符的车型识别方法,其特征在于,所述步骤 1 中对用于训练的车辆视频,跟踪视频中每一种车型的图像,提取该车型每一帧图像的 SIFT 特征,包括以下几个步骤:

步骤 1.1:跟踪训练视频中每一种车型图像;

步骤 1.2:对跟踪到的车型的每一帧图像进行数据增强,得到每种车型的图像数据集;

步骤 1.3:提取每种车型的图像数据集的 SIFT 特征。

3. 根据权利要求 1 所述基于视频费舍尔向量描述符的车型识别方法,其特征在于,所述步骤 2 中对提取的车型图像的所有 SIFT 特征进行费舍尔向量编码计算,包括以下几个步骤:

步骤 2.1:对提取的车型图像数据集的所有 SIFT 特征进行费舍尔向量编码计算,其中车型图像数据集中的一帧图像 1 的费舍尔编码向量表示为  $\phi(1)$ ;

步骤 2.2:对每帧图像 1 的费舍尔编码向量  $\phi(1)$  进行符号平方根处理:

$$\text{sign}(\phi(l)) / \sqrt{|\phi(l)|},$$

再对符号平方根处理后的向量进行 L2 归一化。

4. 根据权利要求 1 所述基于视频费舍尔向量描述符的车型识别方法,其特征在于,所述步骤 3 对编码得到的费舍尔向量描述符进行 PCA 降维,包括学习一个投影矩阵 W,将费舍尔向量描述符从  $R^D$  维降至  $R^M$  维。

5. 根据权利要求 1 所述基于视频费舍尔向量描述符的车型识别方法,其特征在于,所述步骤 4 对降维后的费舍尔向量描述符进行二值化处理,得到该车型的视频费舍尔向量描述符,包括以下几个步骤:

步骤 4.1:步骤 3 中计算得到的费舍尔向量描述符,构成一个矩阵:  $U \in R^{q \times m}$ ;

步骤 4.2:通过符号函数,  $\text{sign}(a) = 1, \text{iff } a > 0$  进行二值化,得到一个二值化向量:  $\beta = \text{sign}(U\psi - w)$

其中  $\psi$  是降维后的费舍尔向量描述符,  $w$  是所选阈值,当  $U\psi > w$ ,  $\beta = 1$ ,反之  $\beta = 0$ ;

步骤 4.3:得到一个 q 位的二值化编码描述符  $\beta = \{0, 1\}^q$ ,也即该车型的视频费舍尔

向量描述符。

6. 根据权利要求 1 所述基于视频费舍尔向量描述符的车型识别方法,其特征在于,所述步骤 5 将得到的所有描述符进行 SVM 训练,得到一个具有 N 个车型类别的识别系统,包括使用 one vs rest SVM 训练一个多类分类器,得到具有 N 类类别的识别系统。

## 一种基于视频费舍尔向量描述符的车型识别方法

### 技术领域

[0001] 本发明属于模式分类技术领域,特别涉及一种基于视频费舍尔向量描述符的车型识别方法。

### 背景技术

[0002] 最近 10 多年来,中国的公路交通基础设施取得了巨大的建设成就,并持续高速的发展着。随着国民经济的快速增长以及中国的城市化进程的持续深化,中国的车辆数量增加迅猛,给环境带来了巨大的压力,也给城市发展和经济增长带来了很多问题。总体上,中国的公路运输事业面临如下挑战:(1) 不断扩大的人口规模和持续快速增长的汽车数量使道路的压力越来越大;(2) 公路交通运输的能耗十分巨大,能源利用不充分;(3) 交通事故高发、交通拥堵状况严重;(4) 大气环境污染十分严重。为了提高交通效率,改善交通状况,现在全世界范围都在兴起使用智能交通系统来管理交通。

[0003] 智能交通系统中的核心功能是对过往车辆的准确检测和正确的车型识别。当前对车辆检测分类技术的研究主要有两个技术流派:车辆自动识别和车辆自动分类。前者是利用车载设备与地面基站设备互识进行,该技术主要用于收费系统中,在发达国家使用范围较广,如美国的 AE-PASS 系统、日本的 ETC 系统,全球卫星 GPS 定位等。后者是通过检测车辆本身固有的参数,在一定车辆分类标准下运用适当的分类识别算法,主动地对车辆进行分型,这一类技术应用比较广泛,已经有很多成熟的系统应用在实际生活中,该类技术可以通过射频微波、红光、激光、声表面波等方式来自动识别车辆信息,也可以使用视频图像处理的方式来识别车牌、车型等车辆信息。比较成熟技术有环形线圈检测、激光红外线检测、超声波/微波检测、地磁检测等,但这几种方法各有优劣,优点是识别精确比较高,但缺点也很明显,主要缺点有施工和安装过程十分复杂,影响正常交通秩序,维护困难,主要设备易损坏,花费较大等。

[0004] 近年来随着计算机多媒体技术和图像处理技术的发展,基于视频的车辆自动分类识别技术在现代交通控制系统中占的分量也越来越大,社会各界投入的研究力量也越来越多。该类技术可以适应动态交通状况的变化,通过实时采集大量的交通流量数据并将其传输到交通管理中心,中心通过系统提供的数据可以迅速做出控制决策,解决交通拥堵等问题。同时,利用该技术可以分析道路的车流量信息,有利于公路网的总体规划及道路建设。视频检测技术的优越性体现在:(1) 采用非接触检测方式,安装维护不必破开路面,不影响路面寿命,不影响交通;(2) 可以检测更大范围内的交通流信息,从而减少设备数量,节约资金;(3) 可以在采集交通流信息的同时提供交通的实时视频图像,便于监察;(4) 对于某些应用,比如交通量调查等,可以把视频图像采集存储后,离线进行分析处理;(5) 当环境发生变化,或系统移动到他处使用时,只需简单设置,系统即可重新投入使用。(6) 可以综合提供交通数据信息和视频图像,便于对现场的全面、直观检测。

[0005] 在图像分类领域,对大规模图像处理一种应用很广的方法是为提取每张照片的局部特征,将提取的特征进行聚类 and 编码得到一个高维向量,然后将其用分类器分类或者使

用最近邻算法进行匹配。其中编码的方法有视觉词袋模型编码,稀疏编码和费舍尔向量编码等。而费舍尔向量编码的性能较其他几种编码方式要好,结合了视觉词袋模型编码和统计模型的优点,能够降低计算量,减小内存消耗,且搜索精度较高。

[0006] 在此基础之上,本专利结合费舍尔向量编码和视频检测技术的优点,提出一种基于视频费舍尔向量描述符的车型识别方法来解决上述问题。本发明能够在视频中进行搜索的车辆图像的车型识别,且识别准确率高,运行速度快,消耗内存少,具有较高的实用性和鲁棒性。

## 发明内容

[0007] 本发明中所述方法是为了克服上述现有技术的缺点,主要针对已经检测到的车辆视频,提取其车辆特征对车型进行细分识别的问题,提出了一种基于视频费舍尔向量描述符的车型识别方法。具体的技术方案如下所述。

[0008] 为解决上述技术问题,本发明采用以下技术方案:

[0009] 一种基于视频费舍尔向量描述符的车型识别方法,包括以下步骤:

[0010] 步骤 1:对用于训练的车辆视频,跟踪视频中每一种车型的图像,提取该车型每一帧图像的 SIFT 特征;

[0011] 步骤 2:对提取的车型图像的所有 SIFT 特征进行费舍尔向量编码计算,得到费舍尔向量;

[0012] 步骤 3:对得到的费舍尔向量描述符进行降维;

[0013] 步骤 4:对降维后的得到的费舍尔向量描述符进行二值化处理,得到该车型的视频费舍尔向量描述符;

[0014] 步骤 5:将得到的所有视频费舍尔向量描述符进行 SVM 训练,得到一个具有 N 个车型类别的识别系统;

[0015] 步骤 6:对测试车辆视频,同样提取视频中车辆图像的视频费舍尔向量描述符,将其导入步骤 5 中训练好的车型识别系统进行测试,识别出测试车辆视频的车型。

[0016] 上述技术方案中,所述步骤 1 中对用于训练的车辆视频,跟踪视频中每一种车型的图像,提取该车型每一帧图像的 SIFT 特征,包括以下几个步骤:

[0017] 步骤 1.1:跟踪训练视频中每一种车型图像;

[0018] 步骤 1.2:对跟踪到的车型的每一帧图像进行数据增强,得到每种车型的图像数据集;

[0019] 步骤 1.3:提取每种车型的图像数据集的 SIFT 特征。

[0020] 上述技术方案中,所述步骤 2 中对提取的车型图像的所有 SIFT 特征进行快速费舍尔向量编码计算,包括以下几个步骤:

[0021] 步骤 2.1:对提取的车型图像数据集的所有 SIFT 特征进行费舍尔向量编码计算,其中车型图像数据集中的一帧图像 1 的费舍尔编码向量表示为  $\phi(1)$ ;

[0022] 步骤 2.2:对每帧图像 1 的费舍尔编码向量  $\phi(1)$  进行符号平方根处理:

[0023] 
$$\text{sign}(\phi(l)) / \sqrt{|\phi(l)|},$$

[0024] 再对符号平方根处理后的向量进行 L2 归一化。

[0025] 上述技术方案中,所述步骤3对编码得到的费舍尔向量描述符进行PCA降维,包括学习一个投影矩阵 $W$ ,将费舍尔向量描述符从 $R^D$ 维降至 $R^m$ 维。

[0026] 上述技术方案中,所述步骤4对降维后的费舍尔向量描述符进行二值化处理,得到该车型的视频费舍尔向量描述符,包括以下几个步骤:

[0027] 步骤4.1:步骤3中计算得到的费舍尔向量描述符,构成一个矩阵: $U \in R^{q \times m}$ ;

[0028] 步骤4.2:通过符号函数,  $\text{sign}(a) = 1, \text{ iff } a > 0$  进行二值化,得到一个二值化向量:

[0029]  $\beta = \text{sign}(U\psi - w)$

[0030] 其中  $\psi$  是降维后的费舍尔向量描述符,  $w$  是所选阈值,当  $U\psi > w$ ,  $\beta = 1$ ,反之  $\beta = 0$ ;

[0031] 步骤4.3:得到一个 $q$ 位的二值化编码描述符  $\beta = \{0, 1\}^q$ ,也即该车型的视频费舍尔向量描述符。

[0032] 上述技术方案中,所述步骤5将得到的所有描述符进行SVM训练,得到一个具有 $N$ 个车型类别的识别系统,包括使用one vs rest SVM训练一个多类分类器,得到具有 $N$ 类类别的识别系统。

[0033] 综上所述,由于采用了上述技术方案,本发明的有益效果是:

[0034] 本发明公开了一种直接对车辆视频处理的车型识别方法,可以快速检测出待识别车辆,采取了费舍尔向量编码来表示视频中的车辆图像,能够降低计算量,减小内存消耗,且搜索精度较高,提高了车型识别率。

#### 附图说明

[0035] 本发明将通过例子并参照附图的方式说明,其中:

[0036] 图1为视频费舍尔向量算法实现示意图。

#### 具体实施方式

[0037] 为详细说明本发明的技术内容、构造特征、所实现目的及效果,以下结合实施方式并配合附图详予说明。

[0038] 本发明提出了一种基于视频费舍尔向量描述符的车型识别方法,在车辆车型识别上取得良好的效果。整个算法实现示意图如图1所示,包括步骤:

[0039] 步骤1:对用于训练的车辆视频,跟踪视频中每一种车型的图像,提取该车型每一帧图像的SIFT特征;

[0040] 主要包括以下几个步骤:

[0041] 步骤1.1:跟踪训练视频中每一种车型图像,由于本发明主要针对于车型识别,在跟踪模块上采取的是基本的粒子滤波技术跟踪视频中的每一种车型的图像,在此不做深入;

[0042] 步骤1.2:对跟踪到的车型的每一帧图像进行数据增强,得到每种车型的图像数据集;

[0043] 步骤1.3:提取每种车型的图像数据集的SIFT特征。

[0044] 目前跟踪技术有很多种,本发明采取基本的粒子滤波技术跟踪视频中的每一种车

型的图像,提取该车型每一帧图像的 SIFT 特征;提取 SIFT 特征包括以下几个步骤:

[0045] 步骤 1.4:尺度空间极值检测:针对构建的高斯金字塔图像,利用差分函数在所有尺度上检测候选极值点。

[0046] 步骤 1.5:关键点定位:通过函数拟合来定位极值点的准确位置和尺度,同时过滤掉候选极值点集合中的边缘点。

[0047] 步骤 1.6:方向确定:结合关键点邻域像素的梯度信息,给每个关键点赋予一个主方向。

[0048] 步骤 1.7:特征描述子生成:对关键点邻域像素的梯度幅值和方向进行统计,得到关键点的特征描述。

[0049] 步骤 2:对提取的车型图像的所有 SIFT 特征进行费舍尔向量编码计算;

[0050] 费舍尔向量编码的本质是利用图像的似然函数的梯度向量来表示一幅图像。一幅图像可由很多局部特征组成,每个特征各个维度之间是独立的,可以用概率模型来模拟这些局部特征,这样,一幅图像的概率分布可表示为各个特征维度上的概率分布的乘积。

[0051] 本专利使用的概率模型是高斯混合模型,根据高斯概率密度函数参数的不同,每个高斯模型看作一个类别,输入一个样本,计算其概率密度函数的值,通过阈值来判断样本是否属于该高斯模型。对于混合高斯模型的费舍尔向量编码,就是把表示图像的高斯分布,分别对其权重、期望和协方差求偏导,用得到的结果表示图像包括以下几个步骤:

[0052] 1.  $X = \{x_t, t = 1 \cdots T\}$  表示为一个图像的描述向量,本专利中  $x_t$  代表每一个 SIFT 描述符,  $\lambda = \{\omega_i, \mu_i, \Sigma_i, i = 1 \cdots N\}$  为参数集,分别表示为模拟图像 SIFT 描述符的高斯函数的权重、期望和协方差矩阵;

[0053] 2. 各 SIFT 特征之间是独立同分布的,则有高斯混合模型的概率密度函数为:

$$[0054] \quad P(X|\lambda) = \prod_{t=1}^T p(x_t|\lambda)$$

[0055] 其中  $p(x_t|\lambda)$  表示高斯混合模型,有:

$$[0056] \quad p(x_t|\lambda) = \sum_{i=1}^N \omega_i p_i(x_t|\lambda)$$

[0057]  $\omega_i$  表示每个单高斯函数的权值,  $p_i(x_t|\lambda)$  表示每个单高斯模型,有:

$$[0058] \quad p_i(x_t|\lambda) = \frac{\exp\left\{-\frac{1}{2}(x_t - \mu_i)^T \Sigma_i^{-1}(x_t - \mu_i)\right\}}{(2\pi)^{D/2} |\Sigma_i|^{1/2}}$$

[0059] 其中 D 表示 SIFT 特征点的维度,高斯混合模型的对数似然函数为:

$$[0060] \quad L(X|\lambda) = \sum_{t=1}^T \log \left\{ \sum_{i=1}^N \omega_i p_i(x_t|\lambda) \right\}$$

[0061] 3. 图像的特征点  $x_t$  由第 i 个高斯分布生成的概率为:

$$[0062] \quad \gamma_i(i) = \frac{\omega_i p_i(x_t|\lambda)}{\sum_{j=1}^N \omega_j p_j(x_t|\lambda)}$$

[0063] 4. 对表示图像的混合高斯分布,分别对其权重、期望和协方差求偏导:

$$[0064] \quad \frac{\partial L(X|\lambda)}{\partial \omega_i} = \sum_{i=1}^T \left[ \frac{\gamma_i(i)}{\omega_i} - \frac{\gamma_i(1)}{\omega_1} \right] \quad (i \geq 2)$$

$$[0065] \quad \frac{\partial L(X|\lambda)}{\partial \mu_i^d} = \sum_{i=1}^T \gamma_i(i) \left[ \frac{x_i^d - \mu_i^d}{(\sigma_i^d)^2} \right]$$

$$[0066] \quad \frac{\partial L(X|\lambda)}{\partial \sigma_i^d} = \sum_{i=1}^T \gamma_i(i) \left[ \frac{(x_i^d - \mu_i^d)^2}{(\sigma_i^d)^3} - \frac{1}{\sigma_i^d} \right]$$

[0067] 其中 d 为特征的维度,  $\sigma$  为高斯分布的标准差;

[0068] 5. 将求偏导后的向量进行标准化, 得到最终的费舍尔编码向量:

$$[0069] \quad f_{\omega_i}^{-1/2} \partial L(X|\lambda) / \partial \omega_i$$

$$[0070] \quad f_{\mu_i^d}^{-1/2} \partial L(X|\lambda) / \partial \mu_i^d$$

$$[0071] \quad f_{\sigma_i^d}^{-1/2} \partial L(X|\lambda) / \partial \sigma_i^d$$

$$[0072] \quad \text{其中 } f_{\omega_i} = T \left( \frac{1}{\omega_i} + \frac{1}{\omega_1} \right), \quad f_{\mu_i^d} = \frac{T \omega_i}{(\sigma_i^d)^2}, \quad f_{\sigma_i^d} = \frac{2T \omega_i}{(\sigma_i^d)^2}。$$

[0073] 步骤 2.1: 对提取的车型图像数据集的所有 SIFT 特征进行费舍尔向量编码计算, 其中车型图像数据集中的一帧图像 1 的费舍尔编码向量表示为  $\phi(1)$ ;

[0074] 步骤 2.2: 对每帧图像 1 的费舍尔编码向量  $\phi(1)$  进行符号平方根处理:

$$[0075] \quad \text{sign}(\phi(l)) / \sqrt{|\phi(l)|},$$

[0076] 再对符号平方根处理后的向量进行 L2 归一化。

[0077] 步骤 3: 对得到的费舍尔向量描述符进行降维;

[0078] 经过步骤 2 得到的费舍尔向量描述符的维数还是很高, 计算的复杂度太大, 通过 PCA 降维, 学习一个投影矩阵 W, 将费舍尔向量描述符降至 R<sup>m</sup> 维。

[0079] 步骤 4: 对降维后的描述符进行二值化处理, 得到该车型的视频费舍尔向量描述符;

[0080] 在步骤 3 降维之后, 为了更进一步降低内存的消耗, 还需要进一步优化, 将降维后的实值向量描述子  $\psi \in R^m$  映射到一个二值化编码  $\beta = \{0, 1\}^q$ , 包括以下几个步骤:

[0081] 步骤 4.1: 步骤 3 中计算得到的费舍尔向量描述符, 构成一个矩阵  $U \in R^{q \times m}$ ;

[0082] 步骤 4.2: 通过符号函数,  $\text{sign}(a) = 1, \text{ iff } a > 0$  进行二值化, 得到一个二值化向量:

$$[0083] \quad \beta = \text{sign}(U\psi - w)$$

[0084] 其中  $\psi$  是降维后的费舍尔向量描述符, w 为所选阈值, 当  $U\psi > w$ ,  $\beta = 1$ , 反之  $\beta = 0$ ;

[0085] 步骤 4.3: 得到一个 q 位的二值化编码描述符  $\beta = \{0, 1\}^q$ , 也即该车型的视频费舍尔向量描述符。



[0086] 步骤5:将得到的所有描述符进行 SVM 训练,得到一个具有 N 个车型类别的识别系统;

[0087] 通过步骤 1 到 4 得到训练车型的视频费舍尔向量描述符,将其进行 one vs rest 线性 SVM 训练,得到一个具有 N 个车型类别的识别系统。

[0088] 步骤6:提取测试车辆视频的视频费舍尔向量描述符,将其导入训练好的 SVM 分类器进行测试,识别出测试车辆视频的车型。

[0089] 对于测试车辆的视频,同样提取其视频费舍尔向量描述符,将其导入之前训练好的识别系统进行识别,包括以下几个步骤:

[0090] 步骤6.1:通过步骤 1 到步骤 4, 提取测试车辆视频的视频费舍尔向量描述符;

[0091] 步骤6.2:将提取的视频费舍尔向量描述符导入训练好的 SVM 分类器进行测试,识别出测试车辆视频的车型。

