

乘客分布预测模型训练方法与系统、乘客诱导方法与系统

技术领域

本发明属于城市轨道交通运营技术领域，尤其涉及一种地铁车厢乘客分布预测模型训练方法与系统，基于环境监测和照明引导的地铁车厢乘客诱导方法与系统。

5 背景技术

随着我国轨道交通系统的不断发展，提升轨道交通系统的运行服务水平成为推动城市轨道交通向前发展的重要力量。地铁乘车过程中由于乘客对车厢各区域的拥挤程度缺乏客观感知，造成部分区域过于拥挤、部分区域闲置，严重影响地铁车厢的空间利用率和乘客乘车体验。因此有必要对车厢和站台的乘客聚集程度进行实时感知，引导乘客向人群稀疏处流动，提高地铁空间利用率。

10

地铁乘客诱导系统及方法包括乘客密度信息获取和诱导信息发布两大关键技术。现阶段乘客密度信息获取主要是采用基于图像检测或压力传感器的方法。基于图像检测的方法在人群流动、遮挡等情况下的乘客密度检测能力无法得到保证；基于压力传感器的方法，加装压力传感器需要对整个地面进行侵入式改造，如公开号为 CN110281979A 的专利申请提出了一种通过在地铁车厢地板安装压电传感器的方法检测车厢内乘客密度。

15

现阶段对乘客密度的诱导系统及方法主要采用语音提示或显示屏提示的方法，如授权公开号为 CN107244330B 的专利提出了一种通过在地铁车厢加装隔离栏、配合语音提示信息引导乘客有序流动，但是语音提示在嘈杂的乘车环境中针对性较差，而显示屏提示难以获得乘客的注意。

20 发明内容

本发明的目的在于提供一种乘客分布预测模型训练方法与系统、乘客诱导方法与系统，以解决在人群流动、遮挡情况下乘客密度检测能力无法得到保证，压电传感器检测乘客密度需要进行侵入式改造，语音提示针对性差以及显示屏提示难以获得乘客的注意等问题。

第一方面，本发明提供一种地铁车厢乘客分布预测模型训练方法，包括：

25

建立每节车厢的空间位置坐标系，在每节车厢安装由温湿度传感器和 CO₂ 浓度传感器构成的采集装置，每个采集装置作为一个采集样本点，获取每个采集样本点在对应空间位置坐标系的安装位置坐标；所述空间位置坐标系是与车顶面或者地面平行的二维直角坐标系，或者所述空间位置坐标系是与车顶面或者地面重合的二维直角坐标系；

获取每节车厢在一个采样周期内每个采集样本点的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列，获取当地当天的平均气温值和平均湿度值，由每个采集样本点的安装位

30

置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂ 浓度时间序列以及平均气温值和平均湿度值构成每节车厢在所述采样周期内每个采集样本点的环境数据样本；

记录所述采样周期的终点时刻每节车厢的乘客位置，并获取乘客在对应空间位置坐标系的位置坐标，即获得每节车厢在所述采样周期内的乘客分布情况；

5 将每节车厢在所述采样周期内的乘客分布情况转化成二值化图像，并计算所述二值化图像中各像素点的均值、方差、极大值和极小值；

计算所述二值化图像中各像素点的幅值和梯度方向，以所有像素点的梯度方向为横轴，以各像素点对应的像素值为纵轴，建立频数分布直方图；并提取所述频数分布直方图的频数分布特征，所述频数分布特征是指将频数分布直方图的横轴等份为 R 个区间，由每个区间
10 对应的像素值构成的特征向量；

将所述像素点的均值、极大值、极小值、方差以及频数分布特征组合构成原始特征向量，以原始特征向量作为聚类算法的输入，对二值化图像进行聚类，得到图像聚类结果；

根据所述图像聚类结果对乘客分布模式进行编码，得到乘客分布编码，每一类图像聚类结果对应一种乘客分布模式，每种乘客分布模式对应一个乘客分布编码；

15 建立每节车厢的分布预测模型，采用训练样本对所述分布预测模型进行训练，得到训练好的分布预测模型，所述训练样本是以每节车厢在所述采样周期内每个采集样本点的环境数据样本作为输入，以对应的乘客分布编码作为输出。

本发明中，利用乘客密度变化引起车厢内温度、湿度和 CO₂ 浓度分布变化的特点，实现乘客分布情况的智能感知，避免了采用图像进行乘客分布检测所面临的人群流动和遮挡问
20 题，以及避免了采用压力传感器进行乘客分布检测所面临的地板侵入式改造难题；将每节车厢内乘客分布数据映射到空间位置坐标系，并对乘客分布情况进行像素化描述，通过图像聚类方式实现了乘客分布模式的智能分类，解决了车厢内乘客分布模式辅助、典型分布模式描述不准的问题。

进一步地，所述空间位置坐标系是以车顶面横向中心线为 x 轴，以车顶面纵向中心线为
25 y 轴的二维直角坐标系，或者所述空间位置坐标系是以地面横向中心线为 x 轴，以地面纵向中心线为 y 轴的二维直角坐标系。

进一步地，在每节车厢的顶部安装 15 个所述采集装置，15 个所述采集装置呈阵列布置。

进一步地，每节车厢在所述采样周期内的乘客分布情况转化成二值化图像的具体过程为：

沿每节车厢所述空间位置坐标系的 x 轴方向将车厢划分为 W 等份，沿每节车厢所述空
30 间位置坐标系的 y 轴方向将车厢划分为 L 等份，将每节车厢空间转化成 $W \times L$ 个独立单元；

设每个所述独立单元为一个像素点,每个所述独立单元内的乘客数量为像素点的像素值;
获取所有像素点的像素值,将乘客分布情况转化成由 $W \times L$ 个像素点组成的二值化图像。
优选的, L 为 3, W 为 10。

进一步地,所述像素点的幅值和梯度方向的计算公式分别为:

$$5 \quad G(l, w) = \sqrt{G_x(l, w)^2 + G_y(l, w)^2}, \quad \theta(l, w) = \arctan \frac{G_y(l, w)}{G_x(l, w)}$$

$$G_x(l, w) = I(l+1, w) - I(l-1, w),$$

$$G_y(l, w) = I(l, w+1) - I(l, w-1)$$

其中, $G(l, w)$ 为像素点的幅值, $\theta(l, w)$ 为像素点的梯度方向, $G_x(l, w)$ 为像素点
沿空间位置坐标系 x 轴方向的梯度幅值, $G_y(l, w)$ 为像素点沿空间位置坐标系 y 轴方向的
10 梯度幅值, $I(l+1, w)$ 为第 $l+1$ 行, 第 w 列像素点的像素值, $I(l-1, w)$ 为第 $l-1$ 行,
第 w 列像素点的像素值, $l=1, 2, 3, \dots, L$, $w=1, 2, 3, \dots, W$, W 为沿每节车厢空间位
置坐标系的 x 轴方向将车厢划分的等份数, L 为沿每节车厢空间位置坐标系的 y 轴方向将车
厢划分的等份数。

进一步地,所述聚类算法为 OPTICS 算法,且设 OPTICS 算法的初始邻域距离参数 ϵ 为
15 0.2, 初始邻域样本参数 $MinPts$ 为 3。

OPTICS 算法对输入参数不敏感,提高了聚类结果的准确度。

优选的,所述分布预测模型为支持向量机模型、相关向量机模型或神经网络模型中的任
一种。

进一步地,将所述训练样本中的温度时间序列、湿度时间序列以及 CO_2 浓度时间序列
20 分别采用温度时间序列特征向量、湿度时间序列特征向量以及 CO_2 浓度时间序列特征向量
来替代,所述温度时间序列特征向量、湿度时间序列特征向量以及 CO_2 浓度时间序列特征
向量的获取方法为:

建立温度时间序列 LSTM (Long Short-Term Memory) 神经网络模型,将训练样本中的
温度时间序列随机分成两个温度训练集,其中一个温度训练集以及与该温度训练集对应的乘
25 客分布编码构成第一训练集,另一个温度训练集为第二训练集;采用第一训练集对温度时间
序列 LSTM 神经网络模型进行训练,得到训练好的温度时间序列 LSTM 神经网络模型;去
除训练好的温度时间序列 LSTM 神经网络模型的分层,以第二训练集作为输入,以全连

接层输出的向量作为输出，得到温度时间序列特征向量；

建立湿度时间序列 LSTM 神经网络模型，将训练样本中的湿度时间序列随机分成两个湿度训练集，其中一个湿度训练集以及与该湿度训练集对应的乘客分布编码构成第三训练集，另一个湿度训练集为第四训练集；采用第三训练集对湿度时间序列 LSTM 神经网络模型进行训练，得到训练好的湿度时间序列 LSTM 神经网络模型；去除训练好的湿度时间序列 LSTM 神经网络模型的分层，以第四训练集作为输入，以全连接层输出的向量作为输出，得到湿度时间序列特征向量；

建立 CO₂ 浓度时间序列 LSTM 神经网络模型，将训练样本中的 CO₂ 浓度时间序列随机分成两个 CO₂ 浓度训练集，其中一个 CO₂ 浓度训练集以及与该 CO₂ 浓度训练集对应的乘客分布编码构成第五训练集，另一个 CO₂ 浓度训练集为第六训练集；采用第五训练集对 CO₂ 浓度时间序列 LSTM 神经网络模型进行训练，得到训练好的 CO₂ 浓度时间序列 LSTM 神经网络模型；去除训练好的 CO₂ 浓度时间序列 LSTM 神经网络模型的分层，以第六训练集作为输入，以全连接层输出的向量作为输出，得到 CO₂ 浓度时间序列特征向量。

原始温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列的维度高，通过 LSTM 神经网络模型提取对应的特征向量，降低了特征向量的维度，降低了训练和测试的复杂度。

第二方面，本发明还提供一种地铁车厢乘客分布预测模型训练系统，包括多个训练子系统，每节车厢对应一个训练子系统，每个所述训练子系统均包括设于车厢顶部的多个采集装置以及控制装置；所述采集装置包括温湿度传感器和 CO₂ 浓度传感器；

所述控制装置包括：
坐标系建立单元，用于建立每节车厢的空间位置坐标系，并获取每个采集样本点在对应空间位置坐标系的安装位置坐标，每个采集装置作为一个采集样本点；

环境数据获取单元，用于获取每节车厢在一个采样周期内每个采集样本点的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列，获取当地当天的平均气温值和平均湿度值，由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂ 浓度时间序列以及平均气温值和平均湿度值构成每节车厢在所述采样周期内每个采集样本点的环境数据样本；

乘客分布获取单元，用于根据记录的所述采样周期终点时刻每节车厢的乘客位置获取乘客在对应空间位置坐标系的位置坐标，即获得每节车厢在所述采样周期内的乘客分布情况；

图像转换及特征提取单元，用于将每节车厢在所述采样周期内的乘客分布情况转化成二值化图像，并计算所述二值化图像中各像素点的均值、方差、极大值和极小值；计算所述二

值化图像中各像素点的幅值和梯度方向，以所有像素点的梯度方向为横轴，以各像素点对应的像素值为纵轴，建立频数分布直方图；并提取所述频数分布直方图的频数分布特征，所述频数分布特征是指将频数分布直方图的横轴等份为 R 个区间，由每个区间对应的像素值构成的特征向量；

- 5 聚类分类单元，用于将所述像素点的均值、极大值、极小值、方差以及频数分布特征组合构成原始特征向量，以原始特征向量作为聚类算法的输入，对二值化图像进行聚类，得到图像聚类结果；

编码单元，用于根据所述图像聚类结果对乘客分布模式进行编码，得到乘客分布编码，每一类图像聚类结果对应一种乘客分布模式，每种乘客分布模式对应一个乘客分布编码；

- 10 模型训练单元，用于建立每节车厢的分布预测模型，采用训练样本对所述分布预测模型进行训练，得到训练好的分布预测模型，所述训练样本是以每节车厢在所述采样周期内每个采集样本点的环境数据样本作为输入，以对应的乘客分布编码作为输出。

第三方面，本发明还提供一种基于环境监测和照明引导的地铁车厢乘客诱导方法，包括：

- 15 实时获取每节车厢在每个采样周期内每个采集样本点的环境数据样本；所述环境数据样本由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、 CO_2 浓度时间序列以及平均气温值和平均湿度值构成；

将所述环境数据样本输入到对应车厢的采用第一方面所述训练方法训练好的分布预测模型中，得到预测的每节车厢在对应采样周期的乘客分布编码；

- 20 根据预测的每节车厢在对应采样周期的乘客分布编码调节对应车厢内照明灯管的亮度，从而引导乘客流动。

采用调整车厢内照明灯管的亮度方式引导乘客流动，例如调暗乘客分布密度大的区域的照明灯管，调亮乘客分布密度小的区域的照明灯管，引导乘客向乘客密度小的区域（即空旷区域）有序流动，避免了车厢嘈杂环境的干扰，并且照明引导方式具有简单、直接、接受度高的优势。

- 25 进一步地，将所述环境数据样本中的温度时间序列、湿度时间序列以及 CO_2 浓度时间序列分别采用温度时间序列特征向量、湿度时间序列特征向量以及 CO_2 浓度时间序列特征向量来替代，所述温度时间序列特征向量、湿度时间序列特征向量以及 CO_2 浓度时间序列特征向量的获取方法为：

- 30 将所述温度时间序列输入至第一方面所述去除分类层的训练好的温度时间序列 LSTM 神经网络模型中，全连接层输出的向量即为对应的温度时间序列特征向量；

将所述湿度时间序列输入至第一方面所述去除分类层的训练好的湿度时间序列 LSTM 神经网络模型中，全连接层输出的向量即为对应的湿度时间序列特征向量；

将所述 CO₂ 浓度时间序列输入至第一方面所述去除分类层的训练好的 CO₂ 浓度时间序列 LSTM 神经网络模型中，全连接层输出的向量即为对应的 CO₂ 浓度时间序列特征向量。

5 进一步地，根据预测的每节车厢在对应采样周期的乘客分布编码控制对应车厢内照明灯管的亮度的具体过程为：

根据预测的每节车厢在对应采样周期的乘客分布编码得到当前图像聚类结果，根据当前图像聚类结果确定当前图像对应的聚类中心；

10 将距离所述聚类中心最近的 S 张二值化图像定义为对应类别的典型聚类图像，计算 S 张所述典型聚类图像中每个像素点的像素平均值，得到各该类别下的像素平均图像；

根据每节车厢照明灯管的布局，为每个照明灯管划分一个照明辖区，并计算每个照明灯管对应照明辖区内像素点的像素平均值；

如果所述像素平均图像中的像素最大值大于像素阈值，则启动照明引导，进入下一个步骤，否则无需照明引导；

15 提取所述像素平均图像中像素平均值大于像素阈值的像素点，由大于像素阈值的像素点构成待引导像素点集合；

根据照明灯管在空间位置坐标系中的坐标、待引导像素点集合中像素点在空间位置坐标系中的坐标以及照明辖区内像素点的像素平均值计算照明引导系数，具体计算表达式为：

$$\beta_n = A_n \left(\sqrt{(x_0^n - x_1^l)^2 + (y_0^n - y_1^w)^2} + \sqrt{(x_0^n - x_2^l)^2 + (y_0^n - y_2^w)^2} + \dots \right. \\ \left. \dots + \sqrt{(x_0^n - x_H^l)^2 + (y_0^n - y_H^w)^2} \right)$$

20 其中， β_n 为照明引导系数， A_n 为第 n 个照明灯管所对应照明辖区内像素点的像素平均值， (x_0^n, y_0^n) 为第 n 个照明灯管在空间位置坐标系的坐标， (x_1^l, y_1^w) 为待引导像素点集合中第 1 个像素点在空间位置坐标系的坐标， (x_2^l, y_2^w) 为待引导像素点集合中第 2 个像素点在空间位置坐标系的坐标， (x_H^l, y_H^w) 为待引导像素点集合中第 H 个像素点在空间位置坐标系的坐标， H 为待引导像素点集合中像素点的数量；

25 按照 β_n 值的排序依次调节与 β_n 值对应的照明辖区内照明灯管的亮度，引导乘客向 β_n 值较小的照明辖区流动。

优选的，像素阈值为 15。

第四方面，本发明还提供一种基于环境监测和照明引导的地铁车厢乘客诱导系统，包括：

实时数据获取单元，用于实时获取每节车厢在每个采样周期内每个采集样本点的环境数据样本；所述环境数据样本由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂ 浓度时间序列以及平均气温值和平均湿度值构成；

预测单元，用于将所述环境数据样本输入到对应车厢的采用第一方面所述训练方法训练好的分布预测模型中，得到预测的每节车厢在对应采样周期的乘客分布编码；

调节及引导单元，用于根据预测的每节车厢在对应采样周期的乘客分布编码调节对应车厢内照明灯管的亮度，从而引导乘客流动。

10 进一步地，所述调节及引导单元，具体用于：

根据预测的每节车厢在对应采样周期的乘客分布编码得到当前图像聚类结果，根据当前图像聚类结果确定当前图像对应的聚类中心；

将距离所述聚类中心最近的 S 张二值化图像定义为对应类别的典型聚类图像，计算 S 张所述典型聚类图像中每个像素点的像素平均值，得到各该类别下的像素平均图像；

15 根据每节车厢照明灯管的布局，为每个照明灯管划分一个照明辖区，并计算每个照明灯管对应照明辖区内像素点的像素平均值；

如果所述像素平均图像中的像素最大值大于像素阈值，则启动照明引导，进入下一个步骤，否则无需照明引导；

20 提取所述像素平均图像中像素平均值大于像素阈值的像素点，由大于像素阈值的像素点构成待引导像素点集合；

根据照明灯管在空间位置坐标系中的坐标、待引导像素点集合中像素点在空间位置坐标系中的坐标以及照明辖区内像素点的像素平均值计算照明引导系数，具体计算表达式为：

$$\beta_n = A_n \left(\sqrt{(x_0^n - x_1^l)^2 + (y_0^n - y_1^w)^2} + \sqrt{(x_0^n - x_2^l)^2 + (y_0^n - y_2^w)^2} + \dots \right. \\ \left. \dots + \sqrt{(x_0^n - x_H^l)^2 + (y_0^n - y_H^w)^2} \right)$$

其中， β_n 为照明引导系数， A_n 为第 n 个照明灯管所对应照明辖区内像素点的像素平均值， (x_0^n, y_0^n) 为第 n 个照明灯管在空间位置坐标系的坐标， (x_1^l, y_1^w) 为待引导像素点集合中第 1 个像素点在空间位置坐标系的坐标， (x_2^l, y_2^w) 为待引导像素点集合中第 2 个像素

点在空间位置坐标系的坐标, (x_H^l, y_H^w) 为待引导像素点集合中第 H 个像素点在空间位置坐标系的坐标, H 为待引导像素点集合中像素点的数量;

按照 β_n 值的排序依次调节与 β_n 值对应的照明辖区内照明灯管的亮度, 引导乘客向 β_n 值较小的照明辖区流动。

5 有益效果

与现有技术相比, 本发明所提供的一种乘客分布预测模型训练方法与系统、乘客诱导方法与系统, 利用乘客密度变化引起车厢内温度、湿度和 CO_2 浓度分布变化的特点, 实现乘客分布情况的智能感知, 避免了采用图像进行乘客分布检测所面临的人群流动和遮挡问题, 以及避免了采用压力传感器进行乘客分布检测所面临的地板侵入式改造难题; 将每节车厢内乘客分布数据映射到空间位置坐标系, 并对乘客分布情况进行像素化描述, 通过图像聚类方式实现了乘客分布模式的智能分类, 解决了车厢内乘客分布模式辅助、典型分布模式描述不准的问题; 采用调整车厢内照明灯管的亮度方式引导乘客流动, 例如调暗乘客分布密度大的区域的照明灯管, 调亮乘客分布密度小的区域的照明灯管, 引导乘客向乘客密度小的区域(即空旷区域)有序流动, 避免了车厢嘈杂环境的干扰, 并且照明引导方式具有简单、直接、接受度高的优势。

附图说明

为了更清楚地说明本发明的技术方案, 下面将对实施例描述中所需要使用的附图作简单地介绍, 显而易见地, 下面描述中的附图仅仅是本发明的一个实施例, 对于本领域普通技术人员来说, 在不付出创造性劳动的前提下, 还可以根据这些附图获得其他的附图。

- 20 图 1 是本发明实施例中一种地铁车厢乘客分布预测模型训练方法流程图;
图 2 是本发明实施例中一种基于环境监测和照明引导的地铁车厢乘客诱导方法流程图;
图 3 是本发明实施例中照明灯管布局及照明辖区示意图。

具体实施方式

下面结合本发明实施例中的附图, 对本发明中的技术方案进行清楚、完整地描述, 显然, 所描述的实施例仅仅是本发明一部分实施例, 而不是全部的实施例。基于本发明中的实施例, 本领域普通技术人员在没有作出创造性劳动的前提下所获得的所有其他实施例, 都属于本发明保护的范围。

如图 1 所示, 本实施例所提供的一种地铁车厢乘客分布预测模型训练方法, 包括:

- 1、获取训练样本
30 1.1 建立坐标系、设置采集装置

建立每节车厢的空间位置坐标系，在每节车厢安装由温湿度传感器和 CO₂ 浓度传感器构成的采集装置，每个采集装置作为一个采集样本点，获取每个采集样本点在对应空间位置坐标系的安装位置坐标。空间位置坐标系是与车顶面或者地面平行的二维直角坐标系，或者空间位置坐标系是与车顶面或者地面重合的二维直角坐标系。空间位置坐标系便于将采集样本点、乘客（或者由乘客构成的独立单元）映射到该空间位置坐标系中。

本实施例中，空间位置坐标系是以车顶面横向中心线为 x 轴，以车顶面纵向中心线为 y 轴的二维直角坐标系，或者空间位置坐标系是以地面横向中心线为 x 轴，以地面纵向中心线为 y 轴的二维直角坐标系。

为了获取地铁车厢内乘客分布与环境参数的映射关系，先采集原始环境数据，原始环境数据由采集装置来采集。在每节车厢的顶部安装 15 个采集装置，15 个采集装置呈 3×5 的阵列布置，设第 i 个采集样本点在对应空间位置坐标系的安装位置坐标为 $P_i = (x_i, y_i)$ ， $i = 1, 2, \dots, o$ ，本实施例中， o 为 15。

1.2 获取训练输入样本（即原始环境数据样本）

获取每节车厢在一个采样周期内每个采集样本点的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列，获取当地当天的平均气温值和平均湿度值，由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂ 浓度时间序列以及平均气温值和平均湿度值构成每节车厢在所述采样周期内每个采集样本点的环境数据样本。

定义地铁到站、屏蔽门开启为采样周期的时间起点 0 ，定义乘客完成上车且位置基本稳定（即无人流动）为采样周期的终点 T ，通过采集装置获取一个采样周期内第 i 个采集样本点的温度时间序列 $Temp^i = \{temp_t^i, t = 1, 2, \dots, \frac{T}{t_0}\}$ 、湿度时间序列 $RH^i = \{Rh_t^i, t = 1, 2, \dots, \frac{T}{t_0}\}$ 和 CO₂ 浓度时间序列 $C_{co2}^i = \{c_t^i, t = 1, 2, \dots, \frac{T}{t_0}\}$ ，其中， t_0 为采样时间间隔， T 为采样周期时长， $temp_t^i$ 为第 i 个采集样本点在时刻 t 的温度值， Rh_t^i 为第 i 个采集样本点在时刻 t 的湿度值， c_t^i 为第 i 个采集样本点在时刻 t 的 CO₂ 浓度值。获取当地当天的平均气温值 $temp_0$ ，平均湿度值 Rh_0 ，则一个采样周期内（时刻 $0 \sim$ 时刻 T ）获取的环境数据样本为 $E = [x_i, y_i, temp_t^i, Rh_t^i, c_t^i, temp_0, Rh_0]$ 。

1.3 获取训练输出样本（即乘客分布编码）

1.31 获取乘客分布情况

为了避免其他乘客分布采集方式因遮挡等导致乘客分布检测准确度低的问题,采用人工统计方式记录采样周期的终点时刻 T 每节车厢的乘客位置,并获取乘客在对应空间位置坐标系的位置坐标,即获得每节车厢在采样周期内的乘客分布情况,具体表示为

$$D = \{x_j, y_j, j = 1, 2, \dots, N\}, \text{ 其中, } j \text{ 为第 } j \text{ 个乘客, } N \text{ 为该节车厢内乘客的数量,}$$

5 (x_j, y_j) 为第 j 个乘客在对应空间位置坐标系的位置坐标。

1.32 每节车厢乘客分布模式编码

为了对乘客分布情况进行定量描述,对乘客分布模式进行编码,具体编码步骤为:

a . 将每节车厢在采样周期内(时刻 0 ~时刻 T)的乘客分布情况转化成二值化图像,具体为:

10 $a1$. 沿每节车厢空间位置坐标系的 x 轴方向将车厢划分为 W 等份,沿每节车厢空间位置坐标系的 y 轴方向将车厢划分为 L 等份,将每节车厢空间转化成 $W \times L$ 个独立单元。本实施例中, $L = 3$, $W = 10$ 。

$a2$. 设每个独立单元为一个像素点,每个独立单元内的乘客数量为像素点的像素值,即第 l 行第 w 列独立单元内的乘客数量 $I(l, w)$ 为第 l 行第 w 列像素点的像素值,

15 $l=1, 2, 3, \dots, L$, $w=1, 2, 3, \dots, W$ 。

$a3$. 获取所有像素点的像素值,将乘客分布情况转化成由 $W \times L$ 个像素点组成的二值化图像。

b . 计算二值化图像中各像素点的均值、方差、极大值 $I_{\max}$ 和极小值 $I_{\min}$ 。

$$\text{二值化图像中各像素点的均值 } \bar{I} \text{ 为 } \bar{I} = \frac{1}{L \times W} \sum_{l=1, w=1}^{l=L, w=W} I(l, w) \quad (1)$$

$$20 \quad \text{二值化图像中各像素点的方差 } I_{SDE} \text{ 为 } I_{SDE} = \sqrt{\sum_{l=1, w=1}^{l=L, w=W} (I(l, w) - \bar{I})^2 / (L \times W)} \quad (2)$$

c . 为了获取二值化图像像素值沿不同方向变化的特征,计算二值化图像中各像素点的幅值和梯度方向,以所有像素点的梯度方向为横轴,以各像素点对应的像素值为纵轴,建立频数分布直方图;并提取频数分布直方图的频数分布特征,频数分布特征是指将频数分布直方图的横轴等份为 R 个区间,由每个区间对应的像素值构成的特征向量。

25 二值化图像中各像素点的幅值和梯度方向的计算公式分别为:

$$G(l, w) = \sqrt{G_x(l, w)^2 + G_y(l, w)^2} \quad (3)$$

$$\theta(l, w) = \arctan \frac{G_y(l, w)}{G_x(l, w)} \quad (4)$$

$$G_x(l, w) = I(l+1, w) - I(l-1, w) \quad (5)$$

$$G_y(l, w) = I(l, w+1) - I(l, w-1) \quad (6)$$

5 其中, $G(l, w)$ 为像素点的幅值, $\theta(l, w)$ 为像素点的梯度方向, $G_x(l, w)$ 为像素点沿空间位置坐标系 x 轴方向的梯度幅值, $G_y(l, w)$ 为像素点沿空间位置坐标系 y 轴方向的梯度幅值, $I(l+1, w)$ 为第 $l+1$ 行, 第 w 列像素点的像素值, $I(l-1, w)$ 为第 $l-1$ 行, 第 w 列像素点的像素值。

本实施例中, $R=9$, 即将梯度方向在 $[0, 180^\circ]$ 内分成 9 等份, 构建由 $[0, 20^\circ)$ 、
10 $[20, 40^\circ)$ 、 $[40, 60^\circ)$ 、 $[60, 80^\circ)$ 、 $[80, 100^\circ)$ 、 $[100, 120^\circ)$ 、 $[120, 140^\circ)$ 、 $[140, 160^\circ)$ 、 $[160, 180^\circ]$ 9 个梯度方向区间。在每个采样周期的时间起点, 各区间的初始频数值为 0, 即梯度方向在该区间的累计像素值为 0, 则初始梯度方向频数向量 $F_0 = \{0, 0, 0, 0, 0, 0, 0, 0, 0\}$, F_0 表示 9 个区间的初始频数值均为 0。设梯度方向频数向量 $F_{l \times w} = \{f_1, \dots, f_r, \dots, f_9\}$, 如果第 $l \times w$ 个像素点的梯度方向落入第 r 个区间, 则更新后的梯度方向频数向量 $F_{l \times w} = \{f_1, \dots, f_r + G(l, w), \dots, f_9\}$, 其中, $l \times w = 1, 2, \dots, L \times W$,
15 $r = 1, 2, \dots, 9$ 。完成所有像素点遍历后获得的最终梯度方向频数向量 $F_{L \times W}$ 为该二值化图像的频数分布特征。

d. 将像素点的均值、极大值、极小值、方差以及频数分布特征组合构成原始特征向量 $P = [\bar{I}, I_{\max}, I_{\min}, I_{SDE}, F_{L \times W}]$, 以原始特征向量作 P 为聚类算法的输入, 对二值化图像
20 进行聚类, 得到图像聚类结果。

采用聚类算法进行分类为现有技术。本实施例中, 聚类算法采用 OPTICS 算法对二值化图像进行聚类, 获得图像聚类结果, 设 OPTICS 算法的初始邻域距离参数 ε 为 0.2, 初始邻域样本参数 $MinPts$ 为 3。OPTICS 算法为现有技术, OPTICS 算法对输入参数不敏感, 提高了聚类结果的准确度。

e. 根据图像聚类结果对乘客分布模式进行编码, 得到乘客分布编码, 每一类图像聚类结果对应一种乘客分布模式, 每种乘客分布模式对应一个乘客分布编码。

聚类结果同属一类的图像为同类图像样本, 同类图像样本对应的乘客分布模式为同类分布模式, 每类乘客分布模式对应为一个乘客分布编码。例如图像聚类结果为将图像分为 5

5 类, 则按照二进制编码可以获得 $[1, 0, 0, 0, 0]^T$, $[0, 1, 0, 0, 0]^T$, $[0, 0, 1, 0, 0]^T$, $[0, 0, 0, 1, 0]^T$ 和 $[0, 0, 0, 0, 1]^T$ 5 个乘客分布编码。按照所属类别中图像样本的数量降序排序, 第 1 类乘客分布编码为 $[1, 0, 0, 0, 0]^T$, 第 2 类乘客分布编码为 $[0, 1, 0, 0, 0]^T$, 第 3 类乘客分布编码为 $[0, 0, 1, 0, 0]^T$, 第 4 类乘客分布编码为 $[0, 0, 0, 1, 0]^T$, 第 5 类乘客分布编码为 $[0, 0, 0, 0, 1]^T$ 。

10 1.4 对原始环境数据样本进行处理

原始环境数据样本中温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列的维度高, 为了降低输入训练样本的维度, 降低后续模型训练和测试的复杂度, 通过建立 LSTM 神经网络模型来提取特征向量, 具体过程为:

建立温度时间序列 LSTM 神经网络模型, 采用第一训练集对温度时间序列 LSTM 神经网络模型进行训练, 得到训练好的温度时间序列 LSTM 神经网络模型; 去除训练好的温度时间序列 LSTM 神经网络模型的分层, 以第二训练集作为输入, 以全连接层输出的向量作为输出, 得到温度时间序列特征向量; 将步骤 1.21 中一个采样周期内的温度时间序列随机分成两个温度训练集, 其中一个温度训练集以及与该温度训练集对应的乘客分布编码(步骤 *e*) 构成第一训练集, 另一个温度训练集为第二训练集。例如, 将一个采样周期内的温度

20 时间序列 $Temp^i = \{temp_t^i, t = 1, 2, \dots, \frac{T}{t_0}\}$ 按照 5:3 分成两个温度训练集, 时间序列长度为 $\frac{T}{t_0}$, 特征维度为采集样本点数量 o 。湿度时间序列和 CO₂ 浓度时间序列类似。

建立湿度时间序列 LSTM 神经网络模型, 采用第三训练集对湿度时间序列 LSTM 神经网络模型进行训练, 得到训练好的湿度时间序列 LSTM 神经网络模型; 去除训练好的湿度时间序列 LSTM 神经网络模型的分层, 以第四训练集作为输入, 以全连接层输出的向量

25 作为输出, 得到湿度时间序列特征向量; 将步骤 1.21 中一个采样周期内的湿度度时间序列随机分成两个湿度训练集, 其中一个湿度训练集以及与该湿度训练集对应的乘客分布编码构成第三训练集, 另一个湿度训练集为第四训练集。

建立 CO₂ 浓度时间序列 LSTM 神经网络模型，采用第五训练集对 CO₂ 浓度时间序列 LSTM 神经网络模型进行训练，得到训练好的 CO₂ 浓度时间序列 LSTM 神经网络模型；去除训练好的 CO₂ 浓度时间序列 LSTM 神经网络模型的分层，以第六训练集作为输入，以全连接层输出的向量作为输出，得到 CO₂ 浓度时间序列特征向量；将步骤 1.21 中一个采样周期内的 CO₂ 浓度时间序列随机分成两个 CO₂ 浓度训练集，其中一个 CO₂ 浓度训练集以及与该 CO₂ 浓度训练集对应的乘客分布编码构成第五训练集，另一个 CO₂ 浓度训练集为第六训练集。

将训练输入样本中的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列分别采用温度时间序列特征向量、湿度时间序列特征向量以及 CO₂ 浓度时间序列特征向量来替代，获得训练输入样本中新的环境数据样本 E' 。

LSTM 神经网络模型的建立为现有技术。本实施例中，LSTM 神经网络模型隐含层神经元数量为 100，训练迭代次数为 2000。2、建立每节车厢的分布预测模型，并对分布预测模型进行训练（即获取环境数据与乘客分布之间的映射关系）

建立每节车厢的分布预测模型，采用训练样本对分布预测模型进行训练，得到训练好的分布预测模型，训练样本是以步骤 1 中该节车厢在一个采样周期内每个采集样本点的新的环境数据样本 E' 作为输入，以对应的乘客分布编码作为输出。

分布预测模型可以为支持向量机模型、相关向量机模型或神经网络模型中的任一种，分布预测模型的建立为现有技术。

本实施例所提供的一种地铁车厢乘客分布预测模型训练系统，包括多个训练子系统，每节车厢对应一个训练子系统，每个训练子系统均包括设于车厢顶部的多个采集装置以及控制装置；采集装置包括温湿度传感器和 CO₂ 浓度传感器。采集装置的设置见步骤 1.1 中的记载。

控制装置可以是整列车共用一控制装置，也可以是每节车厢对应一控制装置，与地铁车厢乘客分布预测模型训练方法对应，控制装置包括：

坐标系建立单元，用于建立每节车厢的空间位置坐标系，并获取每个采集样本点在对应空间位置坐标系的安装位置坐标，每个采集装置作为一个采集样本点；

环境数据获取单元，用于获取每节车厢在一个采样周期内每个采集样本点的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列，获取当地当天的平均气温值和平均湿度值，由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂ 浓度时间序列以及平均气温值和平均湿度值构成每节车厢在采样周期内每个采集样本点的环境数据样本；

乘客分布获取单元,用于根据记录的采样周期终点时刻每节车厢的乘客位置获取乘客在对应空间位置坐标系的位置坐标,即获得每节车厢在采样周期内的乘客分布情况;

图像转换及特征提取单元,用于将每节车厢在采样周期内的乘客分布情况转化成二值化图像,并根据式(1)和(2)计算二值化图像中各像素点的均值、方差、极大值和极小值;
5 根据式(3)~(6)计算二值化图像中各像素点的幅值和梯度方向,以所有像素点的梯度方向为横轴,以各像素点对应的像素值为纵轴,建立频数分布直方图;并提取频数分布直方图的频数分布特征,频数分布特征是指将频数分布直方图的横轴等份为 R 个区间,由每个区间对应的像素值构成的特征向量;

聚类分类单元,用于将像素点的均值、极大值、极小值、方差以及频数分布特征组合构成原始特征向量,以原始特征向量作为聚类算法的输入,对二值化图像进行聚类,得到图像聚类结果;

编码单元,用于根据图像聚类结果对乘客分布模式进行编码,得到乘客分布编码,每一类图像聚类结果对应一种乘客分布模式,每种乘客分布模式对应一个乘客分布编码;

模型训练单元,用于建立每节车厢的分布预测模型,采用训练样本对分布预测模型进行
15 训练,得到训练好的分布预测模型,训练样本是以每节车厢在所述采样周期内每个采集样本点的环境数据样本作为输入,以对应的乘客分布编码作为输出。

如图2所示,本实施例所提供的一种基于环境监测和照明引导的地铁车厢乘客诱导方法,包括:3.1 预测样本的获取

3.11 在预测时,通过步骤1.1设置的采集装置实时获取每节车厢在每个采样周期内每个
20 采集样本点的环境数据样本。环境数据样本由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂浓度时间序列以及平均气温值和平均湿度值构成。

3.12 将步骤3.11的环境数据样本中的温度时间序列输入至步骤1.4中去除分类层的训练好的温度时间序列LSTM神经网络模型中,全连接层输出的向量即为对应的温度时间序列
25 特征向量;

将步骤3.11的环境数据样本中的湿度时间序列输入至步骤1.4中去除分类层的训练好的湿度时间序列LSTM神经网络模型中,全连接层输出的向量即为对应的湿度时间序列特征向量;

将步骤3.11的环境数据样本中的CO₂浓度时间序列输入至步骤1.4中去除分类层的训练好的CO₂浓度时间序列LSTM神经网络模型中,全连接层输出的向量即为对应的CO₂浓
30

度时间序列特征向量。

3.13 采用温度时间序列特征向量、湿度时间序列特征向量以及 CO₂ 浓度时间序列特征向量分别来替代实时获取的环境数据样本中的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列，得到新的实时获取的环境数据样本。

5 3.2 获取预测结果

将新的实时获取的环境数据样本输入到对应车厢的采用地铁车厢乘客分布预测模型训练方法训练好的分布预测模型中，得到预测的每节车厢在对应采样周期的乘客分布编码。

3.3 照明引导

10 以采样周期的终点时刻作为时间起点，根据预测的每节车厢在对应采样周期的乘客分布编码调节对应车厢内照明灯管的亮度，从而引导乘客流动，具体过程为：

3.31 根据预测的每节车厢在对应采样周期的乘客分布编码得到当前图像聚类结果（即当前乘客分布的类别），根据当前图像聚类结果确定当前图像对应的聚类中心。图像聚类结果包括聚类中心以及聚类中心周围的样本。

15 3.32 将距离聚类中心最近的 S 张二值化图像定义为对应类别的典型聚类图像，计算 S 张典型聚类图像中每个像素点的像素平均值，得到各该类别下的像素平均图像。

本实施例中， $S=10$ ，10 张典型聚类图像中位置 (l, w) 处像素点的像素平均值

$$\bar{I}(l, w) = \frac{1}{10} \sum_{k=1}^{10} I(l, w)_k, \quad I(l, w)_k \text{ 为第 } k \text{ 张典型聚类图像在位置 } (l, w) \text{ 的像素值。}$$

20 计算 10 张典型聚类图像中每个像素点的像素平均值，即将 10 张典型聚类图像在相同位置（或对应位置）的像素点的像素求平均值，典型聚类图像上有多少个像素点就有多少个像素平均值。

3.33 根据每节车厢照明灯管的布局，按照就近原则为每个照明灯管划分一个照明辖区，如图 3 所示，双线形平行布局、8 根长条形 LED 灯盘照明为地铁列车的典型照明方式，将控制 8 根长条形 LED 灯管亮度、改变客室的亮度分布，引导乘客流动。

25 计算每个照明灯管对应照明辖区内像素点的像素平均值。与相邻两个照明辖区中照明灯管中心距离相等的点的连线构成该相邻两个照明辖区的分界线。

3.34 如果像素平均图像中的像素最大值 $\bar{I}_{\max}$ 大于像素阈值 δ ，则启动照明引导，进入下一个步骤 3.35（即只要有一个像素点的像素平均值大于像素阈值 δ ，则启动引导），否则无需照明引导。本实施例中， $\delta=15$ ，大于 δ 表示为拥挤状态。

3.35 提取像素平均图像中像素平均值大于像素阈值 δ 的像素点, 由大于像素阈值的像素点构成待引导像素点集合 $\bar{I}_A = \{\bar{I}(l_1, w_1), \bar{I}(l_2, w_2), \dots, \bar{I}(l_H, w_H)\}$, $\bar{I}(l_1, w_1)$ 为待引导像素点中第 1 个像素点的像素平均值, $\bar{I}(l_2, w_2)$ 为待引导像素点中第 2 个像素点的像素平均值, $\bar{I}(l_H, w_H)$ 为待引导像素点中第 H 个像素点的像素平均值, H 为待引导像素点集合中像素点的数量, (l_1, w_1) 表示该待引导像素点中第 1 个像素点为第 l_1 行, 第 w_1 列独立单元所对应的像素点, 即该像素点的像素坐标, 其他像素点可以类推。

3.36 根据照明灯管在空间位置坐标系中的坐标、待引导像素点集合中像素点在空间位置坐标系中的坐标以及照明辖区内像素点的像素平均值计算照明引导系数, 具体计算表达式为:

$$\beta_n = A_n \left(\sqrt{(x_0^n - x_1^l)^2 + (y_0^n - y_1^w)^2} + \sqrt{(x_0^n - x_2^l)^2 + (y_0^n - y_2^w)^2} + \dots \right. \\ \left. \dots + \sqrt{(x_0^n - x_H^l)^2 + (y_0^n - y_H^w)^2} \right) \quad (7)$$

10 其中, β_n 为照明引导系数, A_n 为第 n 个照明灯管所对应照明辖区内像素点的像素平均值, (x_0^n, y_0^n) 为第 n 个照明灯管在空间位置坐标系的坐标, (x_1^l, y_1^w) 为待引导像素点集合中第 1 个像素点在空间位置坐标系的坐标, (x_2^l, y_2^w) 为待引导像素点集合中第 2 个像素点在空间位置坐标系的坐标, (x_H^l, y_H^w) 为待引导像素点集合中第 H 个像素点在空间位置坐标系的坐标。

15 3.37 按照 β_n 值的排序依次调节与 β_n 值对应的照明辖区内照明灯管的亮度, 引导乘客向 β_n 值较小的照明辖区流动。

例如, 将 β_n 值从大到小排序, 将最大 β_n 值对应的照明辖区内照明灯管调为最暗, 将最小 β_n 值对应的照明辖区内照明灯管调为最亮, 中间 β_n 值按照从大到小依次调节对应照明辖区内照明灯管从暗到亮, 通过改变车厢内亮度分布来引导乘客流动, 这种引导方式具有
20 简单、直接、接受度高的优势。

本实施例还提供一种基于环境监测和照明引导的地铁车厢乘客诱导系统, 包括:

实时数据获取单元, 用于实时获取每节车厢在每个采样周期内每个采集样本点的环境数据样本; 环境数据样本由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、 CO_2 浓度时间序列以及平均气温值和平均湿度值构成;

预测单元,用于将环境数据样本输入到对应车厢的采用地铁车厢乘客分布预测模型训练方法训练好的分布预测模型中,得到预测的每节车厢在对应采样周期的乘客分布编码;

调节及引导单元,用于根据预测的每节车厢在对应采样周期的乘客分布编码调节对应车厢内照明灯管的亮度,从而引导乘客流动。

- 5 以上所揭露的仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或变型,都应涵盖在本发明的保护范围之内。

权利要求

1. 一种地铁车厢乘客分布预测模型训练方法, 其特征在于, 包括:

建立每节车厢的空间位置坐标系, 在每节车厢安装由温湿度传感器和 CO₂ 浓度传感器构成的采集装置, 每个采集装置作为一个采集样本点, 获取每个采集样本点对应空间位置坐标系的安装位置坐标;

获取每节车厢在一个采样周期内每个采集样本点的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列, 获取当地当天的平均气温值和平均湿度值, 由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂ 浓度时间序列以及平均气温值和平均湿度值构成每节车厢在所述采样周期内每个采集样本点的环境数据样本;

记录所述采样周期的终点时刻每节车厢的乘客位置, 并获取乘客在对应空间位置坐标系的位置坐标, 即获得每节车厢在所述采样周期内的乘客分布情况;

将每节车厢在所述采样周期内的乘客分布情况转化成二值化图像, 并计算所述二值化图像中各像素点的均值、方差、极大值和极小值;

计算所述二值化图像中各像素点的幅值和梯度方向, 以所有像素点的梯度方向为横轴, 以各像素点对应的像素值为纵轴, 建立频数分布直方图; 并提取所述频数分布直方图的频数分布特征, 所述频数分布特征是指将频数分布直方图的横轴等份为 R 个区间, 由每个区间对应的像素值构成的特征向量;

将所述像素点的均值、极大值、极小值、方差以及频数分布特征组合构成原始特征向量, 以原始特征向量作为聚类算法的输入, 对二值化图像进行聚类, 得到图像聚类结果;

根据所述图像聚类结果对乘客分布模式进行编码, 得到乘客分布编码, 每一类图像聚类结果对应一种乘客分布模式, 每种乘客分布模式对应一个乘客分布编码;

建立每节车厢的分布预测模型, 采用训练样本对所述分布预测模型进行训练, 得到训练好的分布预测模型, 所述训练样本是以每节车厢在所述采样周期内每个采集样本点的环境数据样本作为输入, 以对应的乘客分布编码作为输出。

2. 如权利要求 1 所述的地铁车厢乘客分布预测模型训练方法, 其特征在于: 所述空间位置坐标系是以车顶面横向中心线为 x 轴, 以车顶面纵向中心线为 y 轴的二维直角坐标系, 或者所述空间位置坐标系是以地面横向中心线为 x 轴, 以地面纵向中心线为 y 轴的二维直角坐标系。

3. 如权利要求 1 所述的地铁车厢乘客分布预测模型训练方法, 其特征在于: 在每节车厢的顶部安装 15 个所述采集装置, 15 个所述采集装置呈阵列布置。

4. 如权利要求 1 所述的地铁车厢乘客分布预测模型训练方法, 其特征在于: 每节车厢在所述采样周期内的乘客分布情况转化成二值化图像的具体过程为:

沿每节车厢所述空间位置坐标系的 x 轴方向将车厢划分为 W 等份, 沿每节车厢所述空间位置坐标系的 y 轴方向将车厢划分为 L 等份, 将每节车厢空间转化成 $W \times L$ 个独立单元; 设每个所述独立单元为一个像素点, 每个所述独立单元内的乘客数量为像素点的像素值; 获取所有像素点的像素值, 将乘客分布情况转化成由 $W \times L$ 个像素点组成的二值化图像。

5. 如权利要求 1 所述的地铁车厢乘客分布预测模型训练方法, 其特征在于: 所述像素点的幅值和梯度方向的计算公式分别为:

$$G(l, w) = \sqrt{G_x(l, w)^2 + G_y(l, w)^2}, \quad \theta(l, w) = \arctan \frac{G_y(l, w)}{G_x(l, w)}$$

$$G_x(l, w) = I(l+1, w) - I(l-1, w),$$

$$G_y(l, w) = I(l, w+1) - I(l, w-1)$$

其中, $G(l, w)$ 为像素点的幅值, $\theta(l, w)$ 为像素点的梯度方向, $G_x(l, w)$ 为像素点沿空间位置坐标系 x 轴方向的梯度幅值, $G_y(l, w)$ 为像素点沿空间位置坐标系 y 轴方向的梯度幅值, $I(l+1, w)$ 为第 $l+1$ 行, 第 w 列像素点的像素值, $I(l-1, w)$ 为第 $l-1$ 行, 第 w 列像素点的像素值, $l=1, 2, 3, \dots, L$, $w=1, 2, 3, \dots, W$, w 为沿每节车厢空间位置坐标系的 x 轴方向将车厢划分的等份数, L 为沿每节车厢空间位置坐标系的 y 轴方向将车厢划分的等份数。

6. 如权利要求 1 所述的地铁车厢乘客分布预测模型训练方法, 其特征在于: 所述聚类算法为 OPTICS 算法, 且设 OPTICS 算法的初始邻域距离参数 ϵ 为 0.2, 初始邻域样本参数 $MinPts$ 为 3。

7. 如权利要求 1~6 中任一项所述的地铁车厢乘客分布预测模型训练方法, 其特征在于: 将所述训练样本中的温度时间序列、湿度时间序列以及 CO_2 浓度时间序列分别采用温度时间序列特征向量、湿度时间序列特征向量以及 CO_2 浓度时间序列特征向量来替代, 所述温度时间序列特征向量、湿度时间序列特征向量以及 CO_2 浓度时间序列特征向量的获取方法为:

建立温度时间序列 LSTM 神经网络模型, 将训练样本中的温度时间序列随机分成两个温度训练集, 其中一个温度训练集以及与该温度训练集对应的乘客分布编码构成第一训练集, 另一个温度训练集为第二训练集; 采用第一训练集对温度时间序列 LSTM 神经网络模型进行训练, 得到训练好的温度时间序列 LSTM 神经网络模型; 去除训练好的温度时间序列 LSTM 神经网络模型的分层, 以第二训练集作为输入, 以全连接层输出的向量作为输出, 得到温度时间序列特征向量;

建立湿度时间序列 LSTM 神经网络模型, 将训练样本中的湿度时间序列随机分成两个湿度训练集, 其中一个湿度训练集以及与该湿度训练集对应的乘客分布编码构成第三训练集, 另一个湿度训练集为第四训练集; 采用第三训练集对湿度时间序列 LSTM 神经网络模型进行训练, 得到训练好的湿度时间序列 LSTM 神经网络模型; 去除训练好的湿度时间序列 LSTM 神经网络模型的分层, 以第四训练集作为输入, 以全连接层输出的向量作为输出, 得到湿度时间序列特征向量;

建立 CO_2 浓度时间序列 LSTM 神经网络模型, 将训练样本中的 CO_2 浓度时间序列随机分成两个 CO_2 浓度训练集, 其中一个 CO_2 浓度训练集以及与该 CO_2 浓度训练集对应的乘客分布编码构成第五训练集, 另一个 CO_2 浓度训练集为第六训练集; 采用第五训练集对 CO_2 浓度时间序列 LSTM 神经网络模型进行训练, 得到训练好的 CO_2 浓度时间序列 LSTM 神经网络模型; 去除训练好的 CO_2 浓度时间序列 LSTM 神经网络模型的分层, 以第六训练集作为输入, 以全连接层输出的向量作为输出, 得到 CO_2 浓度时间序列特征向量。

8. 一种地铁车厢乘客分布预测模型训练系统, 包括多个训练子系统, 每节车厢对应一个训练子系统, 其特征在于: 每个所述训练子系统均包括设于车厢顶部的多个采集装置以及控制装置; 所述采集装置包括温湿度传感器和 CO_2 浓度传感器;

所述控制装置包括:

坐标系建立单元, 用于建立每节车厢的空间位置坐标系, 并获取每个采集样本点对应空间位置坐标系的安装位置坐标, 每个采集装置作为一个采集样本点;

环境数据获取单元, 用于获取每节车厢在一个采样周期内每个采集样本点的温度时间序列、湿度时间序列以及 CO_2 浓度时间序列, 获取当地当天的平均气温值和平均湿度值, 由

每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂浓度时间序列以及平均气温值和平均湿度值构成每节车厢在所述采样周期内每个采集样本点的环境数据样本；

乘客分布获取单元，用于根据记录的所述采样周期终点时刻每节车厢的乘客位置获取乘客在对应空间位置坐标系的位置坐标，即获得每节车厢在所述采样周期内的乘客分布情况；

图像转换及特征提取单元，用于将每节车厢在所述采样周期内的乘客分布情况转化成二值化图像，并计算所述二值化图像中各像素点的均值、方差、极大值和极小值；计算所述二值化图像中各像素点的幅值和梯度方向，以所有像素点的梯度方向为横轴，以各像素点对应的像素值为纵轴，建立频数分布直方图；并提取所述频数分布直方图的频数分布特征，所述频数分布特征是指将频数分布直方图的横轴等份为 R 个区间，由每个区间对应的像素值构成的特征向量；

聚类分类单元，用于将所述像素点的均值、极大值、极小值、方差以及频数分布特征组合构成原始特征向量，以原始特征向量作为聚类算法的输入，对二值化图像进行聚类，得到图像聚类结果；

编码单元，用于根据所述图像聚类结果对乘客分布模式进行编码，得到乘客分布编码，每一类图像聚类结果对应一种乘客分布模式，每种乘客分布模式对应一个乘客分布编码；

模型训练单元，用于建立每节车厢的分布预测模型，采用训练样本对所述分布预测模型进行训练，得到训练好的分布预测模型，所述训练样本是以每节车厢在所述采样周期内每个采集样本点的环境数据样本作为输入，以对应的乘客分布编码作为输出。

9. 一种基于环境监测和照明引导的地铁车厢乘客诱导方法，其特征在于，包括：

实时获取每节车厢在每个采样周期内每个采集样本点的环境数据样本；所述环境数据样本由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、CO₂浓度时间序列以及平均气温值和平均湿度值构成；

将所述环境数据样本输入到对应车厢的采用权利要求 1~7 中任一项所述训练方法训练好的分布预测模型中，得到预测的每节车厢在对应采样周期的乘客分布编码；

根据预测的每节车厢在对应采样周期的乘客分布编码调节对应车厢内照明灯管的亮度，从而引导乘客流动。

10. 如权利要求 9 所述的地铁车厢乘客分布预测模型训练方法，其特征在于：将所述环境数据样本中的温度时间序列、湿度时间序列以及 CO₂ 浓度时间序列分别采用温度时间序列特征向量、湿度时间序列特征向量以及 CO₂ 浓度时间序列特征向量来替代，所述温度时间序列特征向量、湿度时间序列特征向量以及 CO₂ 浓度时间序列特征向量的获取方法为：

将所述温度时间序列输入至权利要求 7 所述去除分类层的训练好的温度时间序列 LSTM 神经网络模型中，全连接层输出的向量即为对应的温度时间序列特征向量；

将所述湿度时间序列输入至权利要求 7 所述去除分类层的训练好的湿度时间序列 LSTM 神经网络模型中，全连接层输出的向量即为对应的湿度时间序列特征向量；

将所述 CO₂ 浓度时间序列输入至权利要求 7 所述去除分类层的训练好的 CO₂ 浓度时间序列 LSTM 神经网络模型中，全连接层输出的向量即为对应的 CO₂ 浓度时间序列特征向量。

11. 如权利要求 9 或 10 所述的地铁车厢乘客诱导方法，其特征在于：根据预测的每节车厢在对应采样周期的乘客分布编码控制对应车厢内照明灯管的亮度的具体过程为：

根据预测的每节车厢在对应采样周期的乘客分布编码得到当前图像聚类结果，根据当前图像聚类结果确定当前图像对应的聚类中心；

将距离所述聚类中心最近的 S 张二值化图像定义为对应类别的典型聚类图像，计算 S 张所述典型聚类图像中每个像素点的像素平均值，得到各该类别下的像素平均图像；

根据每节车厢照明灯管的布局，为每个照明灯管划分一个照明辖区，并计算每个照明灯

管对应照明辖区内像素点的像素平均值；

如果所述像素平均图像中的像素最大值大于像素阈值，则启动照明引导，进入下一个步骤，否则无需照明引导；

提取所述像素平均图像中像素平均值大于像素阈值的像素点，由大于像素阈值的像素点构成待引导像素点集合；

根据照明灯管在空间位置坐标系中的坐标、待引导像素点集合中像素点在空间位置坐标系中的坐标以及照明辖区内像素点的像素平均值计算照明引导系数，具体计算表达式为：

$$\beta_n = A_n \left(\sqrt{(x_0^n - x_1^l)^2 + (y_0^n - y_1^w)^2} + \sqrt{(x_0^n - x_2^l)^2 + (y_0^n - y_2^w)^2} + \dots \right. \\ \left. \dots + \sqrt{(x_0^n - x_H^l)^2 + (y_0^n - y_H^w)^2} \right)$$

其中， β_n 为照明引导系数， A_n 为第 n 个照明灯管所对应照明辖区内像素点的像素平均值， (x_0^n, y_0^n) 为第 n 个照明灯管在空间位置坐标系的坐标， (x_1^l, y_1^w) 为待引导像素点集合中第 1 个像素点在空间位置坐标系的坐标， (x_2^l, y_2^w) 为待引导像素点集合中第 2 个像素点在空间位置坐标系的坐标， (x_H^l, y_H^w) 为待引导像素点集合中第 H 个像素点在空间位置坐标系的坐标， H 为待引导像素点集合中像素点的数量；

按照 β_n 值的排序依次调节与 β_n 值对应的照明辖区内照明灯管的亮度，引导乘客向 β_n 值较小的照明辖区流动。

12. 一种基于环境监测和照明引导的地铁车厢乘客诱导系统，包括：

实时数据获取单元，用于实时获取每节车厢在每个采样周期内每个采集样本点的环境数据样本；所述环境数据样本由每个采集样本点的安装位置坐标、该采集样本点对应的温度时间序列、湿度时间序列、 CO_2 浓度时间序列以及平均气温值和平均湿度值构成；

预测单元，用于将所述环境数据样本输入到对应车厢的采用权利要求 1~7 中任一项所述训练方法训练好的分布预测模型中，得到预测的每节车厢在对应采样周期的乘客分布编码；

调节及引导单元，用于根据预测的每节车厢在对应采样周期的乘客分布编码调节对应车厢内照明灯管的亮度，从而引导乘客流动。

13. 如权利要求 12 所述的地铁车厢乘客诱导系统，其特征在于，所述调节及引导单元，具体用于：

根据预测的每节车厢在对应采样周期的乘客分布编码得到当前图像聚类结果，根据当前图像聚类结果确定当前图像对应的聚类中心；

将距离所述聚类中心最近的 S 张二值化图像定义为对应类别的典型聚类图像，计算 S 张所述典型聚类图像中每个像素点的像素平均值，得到各该类别下的像素平均图像；

根据每节车厢照明灯管的布局，为每个照明灯管划分一个照明辖区，并计算每个照明灯管对应照明辖区内像素点的像素平均值；

如果所述像素平均图像中的像素最大值大于像素阈值，则启动照明引导，进入下一个步骤，否则无需照明引导；

提取所述像素平均图像中像素平均值大于像素阈值的像素点，由大于像素阈值的像素点构成待引导像素点集合；

根据照明灯管在空间位置坐标系中的坐标、待引导像素点集合中像素点在空间位置坐标

系中的坐标以及照明辖区内像素点的像素平均值计算照明引导系数，具体计算表达式为：

$$\beta_n = A_n \left(\sqrt{(x_0^n - x_1^l)^2 + (y_0^n - y_1^w)^2} + \sqrt{(x_0^n - x_2^l)^2 + (y_0^n - y_2^w)^2} + \cdots \right. \\ \left. \cdots + \sqrt{(x_0^n - x_H^l)^2 + (y_0^n - y_H^w)^2} \right)$$

其中， β_n 为照明引导系数， A_n 为第 n 个照明灯管所对应照明辖区内像素点的像素平均值， (x_0^n, y_0^n) 为第 n 个照明灯管在空间位置坐标系的坐标， (x_1^l, y_1^w) 为待引导像素点集合中第1个像素点在空间位置坐标系的坐标， (x_2^l, y_2^w) 为待引导像素点集合中第2个像素点在空间位置坐标系的坐标， (x_H^l, y_H^w) 为待引导像素点集合中第 H 个像素点在空间位置坐标系的坐标， H 为待引导像素点集合中像素点的数量；

按照 β_n 值的排序依次调节与 β_n 值对应的照明辖区内照明灯管的亮度，引导乘客向 β_n 值较小的照明辖区流动。