



(12)发明专利申请

(10)申请公布号 CN 111161322 A

(43)申请公布日 2020.05.15

(21)申请号 201911418564.6

(22)申请日 2019.12.31

(71)申请人 大连理工大学

地址 116024 辽宁省大连市高新园区凌工
路2号

(72)发明人 连静 王欣然 李琳辉 周雅夫
周彬 杨曰凯

(74)专利代理机构 大连东方专利代理有限责任
公司 21212

代理人 李洪福

(51)Int.Cl.

G06T 7/246(2017.01)

G06T 7/207(2017.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

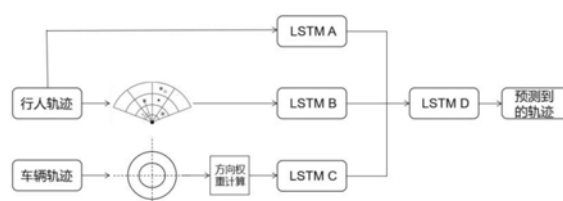
权利要求书3页 说明书4页 附图1页

(54)发明名称

一种基于人车交互的LSTM神经网络行人轨迹预测方法

(57)摘要

本发明公开了一种基于人车交互的LSTM神经网络行人轨迹预测方法,包括以下步骤:构建人车交互的LSTM神经网络;建立多层神经网络的输入,包括输入当前行人轨迹、输入人-人交互信息、输入人-车交互信息;建立人车交互的LSTM神经网络的输出。本发明选用行人的前进方向和速度作为输入,更直观的表现出邻居和车辆对当前行人运动的影响。本发明引入邻居行人和车辆的影响作为“社会信息”,与行人轨迹共同作为输入,运用分层编码的模式构建神经网络,能够解决行人受社会因素影响所发生轨迹改变的问题,提高预测精度。本发明提出方向注意力函数,用来区分不同方向的车辆对行人的影响,以提高“社会信息”的精度,进而提高行人轨迹预测精度。



1. 一种基于人车交互的LSTM神经网络行人轨迹预测方法,其特征在于:包括以下步骤:

A、构建人车交互的LSTM神经网络

所述的人车交互的LSTM神经网络包括行人轨迹输入、车辆轨迹输入、4个LSTM层和预测轨迹层,所述的4个LSTM层分别为LSTM层A、LSTM层B、LSTM层C和LSTM层D,所述的LSTM层A、LSTM层B、LSTM层C的输出端链接后输入到LSTM层D;所述的LSTM层A的输入端接收行人轨迹输入,所述的LSTM层B的输入端经人-人交互栅格地图连接接收行人轨迹输入,所述的LSTM层C的输入端经方向权重计算模块和人-车交互栅格地图接收车辆轨迹输入,所述的LSTM层D的输出端为预测到的轨迹;

所述的LSTM层A对行人轨迹输入进行编码;

所述的LSTM层B对符合人-人交互栅格地图位置条件的行人轨迹输入进行编码;

所述的LSTM层C对方向权重计算模块的输出进行编码,所述的方向权重计算模块对符合人-车交互栅格地图位置条件的车辆轨迹输入进行编码;

所述的LSTM层D对LSTM层A、LSTM层B和LSTM层C链接后的输出进行解码和计算,得到预测的行人轨迹;

B、建立多层神经网络的输入

B1、输入当前行人轨迹

行人行走过程中,若没有其他障碍物阻挡,会沿原来的运动方向一直前进,若遇到其他车辆或行人,会修正自己的运动轨迹,这种修正行为最直观的体现在行人的前进的方向和速度上;行人会依据周边环境,修正自己的方向和速度,进而改变自己未来的轨迹;因此,在行人轨迹预测问题中,使用行人的前进方向和速度作为当前行人轨迹的输入变量;

首先假定每个场景中的行人以当前场景为标准坐标系,定义每个时刻的所有行人坐标,时间以帧为单位;将一段行人时间序列分为观察帧和预测帧,取 $t \sim t+obs$ 为观察帧, $t+obs+1 \sim t+obs+pred$ 为预测帧,其中obs、pred分别为观察帧和预测帧的长度,单位为帧数,则预测问题变成分段预测问题,其输入数据是行人在过去一段时间、每一帧相对于上一帧行人的转向角和行人的速度;具体公式如下:

$$person_{obs}^i = \{(x_t^i, y_t^i), (x_{t+1}^i, y_{t+1}^i) \dots (x_{t+obs}^i, y_{t+obs}^i)\} \quad (1)$$

$$x_{pred}^i = \{(x_{t+obs+1}^i, y_{t+obs+1}^i) \dots (x_{t+obs+pred}^i, y_{t+obs+pred}^i)\} \quad (2)$$

$$v_{obs}^i = \{(vx_t^i, vy_t^i), (vx_{t+1}^i, vy_{t+1}^i) \dots (vx_{t+obs}^i, vy_{t+obs}^i)\} \quad (3)$$

$$angle_{obs}^i = \left\{ \left(\arctan \left(\frac{y_t^i - y_{t-1}^i}{x_t^i - x_{t-1}^i} \right) \right), \arctan \left(\frac{y_{t+1}^i - y_t^i}{x_{t+1}^i - x_t^i} \right) \dots \arctan \left(\frac{y_{t+obs-1}^i - y_{t+obs}^i}{x_{t+obs-1}^i - x_{t+obs}^i} \right) \right\} \quad (4)$$

$$input_t^i = \{angle_{obs}^i, v_{obs}^i\} \quad (5)$$

$$output_t^i = \{angle_{pred}^i, v_{pred}^i\} \quad (6)$$

其中, $person_{obs}^i$ 是观察帧序列中的行人i位置, v_{obs}^i 是观察帧序列中行人i的速度, vx_t^i 是t时刻行人i在x轴方向的速度,单位为m/s; vy_t^i 是t时刻行人i在y轴方向的速度,单位为m/s;

$angle_{obs}^i$ 是观察帧序列中, 行人 i 的方向角, 单位为rad; $input_t^i$ 是行人 i 的输入序列, $output_t^i$ 是行人 i 的输出序列;

其中 x_t^i 、 y_t^i 分别为第 i 个行人在第 t 时刻的 x 、 y 坐标; 将输入值输入到步骤A中的LSTM层A, 输出为:

$$p_t^i = LSTM(p_{t-1}^i, input_t^i; W_1) \quad (7)$$

其中 p_t^i 是LSTM层A的输出, W_1 是LSTM层A的权重;

B2、输入人-人交互信息

建立扇形栅格地图的方式来衡量当前预测行人周围的行人状况, 以当前行人作为圆心, r 为半径的扇形区域为人-人社交感兴趣区域, 在这个区域中的行人称为当前行人的“邻居”, 将扇形区域划分为 $a \times b$ 个网格, 构成栅格地图; 将栅格地图写成 $a \times b$ 的矩阵的形式, 不同网格之内的行人的数量构成 $a \times b$ 的矩阵中的元素, 如下式所示:

$$O_t^i(a, b) = \sum_{j \in N^i} (1_{ab}(s_t^{i,j}, r_t^{i,j})) \quad (8)$$

其中, $1_{ab}(\cdot)$ 是判定邻居行人 j 是否在感兴趣区域中网格 (a, b) 上的判别函数; 若在网格 (a, b) 中则为1; N 是半径为 r 的扇形感兴趣区域, $s_t^{i,j}$ 是预测行人 i 与邻居行人 j 之间的欧氏距离, $r_t^{i,j}$ 是被预测行人 i 与邻居行人 j 位置连线与扇形左侧边的夹角;

将 $O_t^i(a, b)$ 作为输入, 输入到步骤A中建立的LSTM层B中, 则输出为:

$$s_o^{i,t} = LSTM(s_o^{i,t-1}, O_t^i; W_2) \quad (9)$$

$s_o^{i,t}$ 是行人 i 在时刻 t 的LSTM层B的输出, W_2 是LSTM层B的权重;

B3、输入人-车交互信息

建立圆形的栅格地图来表达车辆与行人交互的邻域, 以当前行人作为圆心 R 为半径, 在圆形区域内的车辆为行人关注的车辆, 将圆形地图分为 $m \times n$ 大小的四个网格, 构建行人对车辆的地图矩阵如下:

$$C_t^i(m, n) = \sum_{j \in N^i} 1_{mn}(x_t^j, y_t^j) \quad (10)$$

其中, $1_{mn}(\cdot)$ 是判定车辆 j 是否在感兴趣区域中网格 (m, n) 上的判别函数; 若在网格 (m, n) 中则为1; N 是半径为 R 的圆形兴趣区域, x_t^j 、 y_t^j 是车辆 j 在 t 时刻的横纵坐标;

由于车辆行驶方向相对固定, 行人往往更关注于即将与自己相撞的车辆, 若车辆与行人相对而行, 行人大概率会修正自己的轨迹以避让或加速前进, 若车辆与行人背道而行, 车辆的运动状态对行人的影响相对较小; 为了量化行人与车辆的方向对行人轨迹的影响, 设计方向注意系数 $f_t^{i,j}$ 来计算栅格地图上各点的权重:

$$f_t^{i,j} = c \exp(s_t^{i,j} - s_{t-1}^{i,j}) \quad (11)$$

其中 $s_t^{i,j}$ 、 $s_{t-1}^{i,j}$ 分别是 t 、 $t-1$ 时刻被预测行人 i 与车辆 j 之间的距离, c 是校准方向系数;

因此, 人-车交互输入 $V_t^i(m, n)$ 为:

$$C_t^i(m, n) = \sum_{j \in N^i} (\alpha_{mn}(s_t^{i,j}, r_t^{i,j}) \times f_t^{i,j}) \quad (12)$$

其中, $C_t^i(m, n)$ 是在t时刻、第i个行人对车辆的地图构建的矩阵, $\alpha_{mn}(m, n)$ 用来判断第j个车辆是否在构建的地图内;因此步骤A中LSTM层C的输出为:

$$v_c^{i,t} = LSTM(v_c^{i,t-1}, C_t^i; W_3) \quad (13)$$

$v_c^{i,t}$ 是行人i在t时刻的LSTM层C的输出, W_3 是LSTM层C的权重;

C、建立人车交互的LSTM神经网络的输出

将步骤B中LSTM层A、LSTM层B和LSTM层C的三个输出链接到一起输入到LSTM层D中,输出预测的行人轨迹,如下式所示:

$$h_t^i = p_t^i \oplus s_t^i \oplus v_t^i \quad (14)$$

$$r_c^{i,t} = LSTM(r_c^{i,t-1}, h_t^i; W_4) \quad (15)$$

$$(\hat{x}_t^i, \hat{y}_t^i) = W_o h_t^i + b_o \quad (16)$$

其中 \oplus 表示链接运算符, h_t^i 是LTSM层D的输入, r_t^i 是LTSM层D的输出, W_4 是LTSM层D的权重; \hat{x}_t^i 、 \hat{y}_t^i 是最终预测到的行人轨迹坐标, W_o , b_o 分别是LTSM层D的权重和偏移量。

一种基于人车交互的LSTM神经网络行人轨迹预测方法

技术领域

[0001] 本发明涉及一种行人轨迹预测方法,特别涉及到一种基于人-车交互的LSTM(Long Short-term Memory Networks,简称LSTM)神经网络的预测方法。

背景技术

[0002] 在自动驾驶技术尤其是高密度混合交通环境中的自动驾驶技术的研究中,行人轨迹预测具有重要的意义。混合交通环境中,行人、非机动车与机动车等具备观察、思考、决策和行动能力的“智能单元”在同一道路空间活动,在共享区域内交互影响、互相作用。对行驶在混合交通环境中的智能汽车而言,在交互时仅仅避让是不够的,会导致车辆长久性“踌躇不前”或“进退失据”,并且其新颖性带来的围观行为更有可能引起交通流异常。因此,智能汽车在行驶中对行人在未来一段时间进行精准的轨迹预测具有重要的意义。行人轨迹预测是一个典型的时间序列预测问题,具有高度的时变性和非线性。由于行人轨迹预测本质是一个时间序列问题,随着深度学习的发展,基于数据驱动的建模方式成为研究热点,LSTM神经网络具有RNN(Recurrent Neural Network)网络输出反馈的特点,又能够通过自身内部的cell结构选择性长期记住需要记住的信息,可以更好的对行人轨迹进行预测。但是单纯的LSTM建模的方式忽略了行人与行人、行人与周围环境的交互影响,无法解释行人为了避免碰撞或结伴而行产生的轨迹改变问题。

发明内容

[0003] 针对上述问题,本发明要提出一种能提高行人轨迹预测精度的基于人车交互的LSTM神经网络行人轨迹预测方法,以实现良好的多步预测动态特性。

[0004] 为了实现上述目的,本发明的技术方案如下:一种基于人车交互的LSTM神经网络行人轨迹预测方法,包括以下步骤:

[0005] A、构建人车交互的LSTM神经网络

[0006] 所述的人车交互的LSTM神经网络包括行人轨迹输入、车辆轨迹输入、4个LSTM层和预测轨迹层,所述的4个LSTM层分别为LSTM层A、LSTM层B、LSTM层C和LSTM层D,所述的LSTM层A、LSTM层B、LSTM层C的输出端链接后输入到LSTM层D;所述的LSTM层A的输入端接收行人轨迹输入,所述的LSTM层B的输入端经人-人交互栅格地图连接接收行人轨迹输入,所述的LSTM层C的输入端经方向权重计算模块和人-车交互栅格地图接收车辆轨迹输入,所述的LSTM层D的输出端为预测到的轨迹;

[0007] 所述的LSTM层A对行人轨迹输入进行编码;

[0008] 所述的LSTM层B对符合人-人交互栅格地图位置条件的行人轨迹输入进行编码;

[0009] 所述的LSTM层C对方向权重计算模块的输出进行编码,所述的方向权重计算模块对符合人-车交互栅格地图位置条件的车辆轨迹输入进行编码;

[0010] 所述的LSTM层D对LSTM层A、LSTM层B和LSTM层C链接后的输出进行解码和计算,得到预测的行人轨迹。

[0011] B、建立多层神经网络的输入

[0012] B1、输入当前行人轨迹

[0013] 行人行走过程中,若没有其他障碍物阻挡,会沿原来的运动方向一直前进,若遇到其他车辆或行人,会修正自己的运动轨迹,这种修正行为最直观的体现在行人的前进的方向和速度上。行人会依据周边环境,修正自己的方向和速度,进而改变自己未来的轨迹。因此,在行人轨迹预测问题中,使用行人的前进方向和速度作为当前行人轨迹的输入变量。

[0014] 首先假定每个场景中的行人以当前场景为标准坐标系,定义每个时刻的所有行人坐标,时间以帧为单位。将一段行人时间序列分为观察帧和预测帧,取 $t \sim t+obs$ 为观察帧, $t+obs+1 \sim t+obs+pred$ 为预测帧,其中 obs 、 $pred$ 分别为观察帧和预测帧的长度,单位为帧数,则预测问题变成分段预测问题,其输入数据是行人在过去一段时间、每一帧相对于上一帧行人的转向角和行人的速度。具体公式如下:

$$[0015] \quad person_{obs}^i = \{(x_t^i, y_t^i), (x_{t+1}^i, y_{t+1}^i) \dots (x_{t+obs}^i, y_{t+obs}^i)\} \quad (1)$$

$$[0016] \quad x_{pred}^i = \{(x_{t+obs+1}^i, y_{t+obs+1}^i) \dots (x_{t+obs+pred}^i, y_{t+obs+pred}^i)\} \quad (2)$$

$$[0017] \quad v_{obs}^i = \{(vx_t^i, vy_t^i), (vx_{t+1}^i, vy_{t+1}^i) \dots (vx_{t+obs}^i, vy_{t+obs}^i)\} \quad (3)$$

$$[0018] \quad angle_{obs}^i = \left\{ \left(\arctan \left(\frac{y_t^i - y_{t-1}^i}{x_t^i - x_{t-1}^i} \right) \right), \arctan \left(\frac{y_{t+1}^i - y_t^i}{x_{t+1}^i - x_t^i} \right) \dots \arctan \left(\frac{y_{t+obs-1}^i - y_{t+obs}^i}{x_{t+obs-1}^i - x_{t+obs}^i} \right) \right\} \quad (4)$$

$$[0019] \quad input_t^i = \{angle_{obs}^i, v_{obs}^i\} \quad (5)$$

$$[0020] \quad output_t^i = \{angle_{pred}^i, v_{pred}^i\} \quad (6)$$

[0021] 其中, $person_{obs}^i$ 是观察帧序列中的行人 i 位置, v_{obs}^i 是观察帧序列中行人 i 的速度, vx_t^i 是 t 时刻行人 i 在 x 轴方向的速度,单位为 m/s ; vy_t^i 是 t 时刻行人 i 在 y 轴方向的速度,单位为 m/s ; $angle_{obs}^i$ 是观察帧序列中,行人 i 的方向角,单位为 rad ; $input_t^i$ 是行人 i 的输入序列, $output_t^i$ 是行人 i 的输出序列。

[0022] 其中 x_t^i 、 y_t^i 分别为第 i 个行人在第 t 时刻的 x 、 y 坐标。将输入值输入到步骤A中的LSTM层A,输出为:

$$[0023] \quad p_t^i = LSTM(p_{t-1}^i, input_t^i; W_1) \quad (7)$$

[0024] 其中 p_t^i 是LSTM层A的输出, W_1 是LSTM层A的权重。

[0025] B2、输入人-人交互信息

[0026] 建立扇形栅格地图的方式来衡量当前预测行人周围的行人状况,以当前行人作为圆心, r 为半径的扇形区域为人-人社交感兴趣区域,在这个区域中的行人称为当前行人的“邻居”,将扇形区域划分为 $a \times b$ 个网格,构成栅格地图。将栅格地图写成 $a \times b$ 的矩阵的形式,不同网格之内的行人的数量构成 $a \times b$ 的矩阵中的元素,如下式所示:

$$[0027] \quad O_t^i(a, b) = \sum_{j \in N^i} (1_{ab}(s_t^{i,j}, r_t^{i,j})) \quad (8)$$

[0028] 其中, $l_{ab}(\cdot)$ 是判定邻居行人 j 是否在感兴趣区域中网格 (a, b) 上的判别函数。若在网格 (a, b) 中则为1; N 是半径为 r 的扇形感兴趣区域, $s_t^{i,j}$ 是预测行人 i 与邻居行人 j 之间的欧氏距离, $r_t^{i,j}$ 是被预测行人 i 与邻居行人 j 位置连线与扇形左侧边的夹角。

[0029] 将 $O_t^i(a, b)$ 作为输入, 输入到步骤A中建立的LSTM层B中, 则输出为:

$$[0030] \quad s_o^{i,t} = LSTM(s_o^{i,t-1}, O_t^i; W_2) \quad (9)$$

[0031] $s_o^{i,t}$ 是行人 i 在时刻 t 的LSTM层B的输出, W_2 是LSTM层B的权重。

[0032] B3、输入人-车交互信息

[0033] 建立圆形的栅格地图来表达车辆与行人交互的邻域, 以当前行人作为圆心 R 为半径, 在圆形区域内的车辆为行人关注的车辆, 将圆形地图分为 $m \times n$ 大小的四个网格, 构建行人对车辆的地图矩阵如下:

$$[0034] \quad C_t^i(m, n) = \sum_{j \in N^i} 1_{mn}(x_t^j, y_t^j) \quad (10)$$

[0035] 其中, $1_{mn}(\cdot)$ 是判定车辆 j 是否在感兴趣区域中网格 (m, n) 上的判别函数。若在网格 (m, n) 中则为1; N 是半径为 R 的圆形兴趣区域, x_t^j , y_t^j 是车辆 j 在 t 时刻的横纵坐标。

[0036] 由于车辆行驶方向相对固定, 行人往往更关注于即将与自己相撞的车辆, 若车辆与行人相对而行, 行人大概率会修正自己的轨迹以避免或加速前进, 若车辆与行人背道而行, 车辆的运动状态对行人的影响相对较小。为了量化行人与车辆的方向对行人轨迹的影响, 设计方向注意系数 $f_t^{i,j}$ 来计算栅格地图上各点的权重:

$$[0037] \quad f_t^{i,j} = c \exp(s_t^{i,j} - s_{t-1}^{i,j}) \quad (11)$$

[0038] 其中 $s_t^{i,j}$ 、 $s_{t-1}^{i,j}$ 分别是 t 、 $t-1$ 时刻被预测行人 i 与车辆 j 之间的距离, c 是校准方向系数。

[0039] 因此, 人-车交互输入 $V_t^i(m, n)$ 为:

$$[0040] \quad C_t^i(m, n) = \sum_{j \in N^i} (\alpha_{mn}(s_t^{i,j}, r_t^{i,j}) \times f_t^{i,j}) \quad (12)$$

[0041] 其中, $C_t^i(m, n)$ 是在 t 时刻、第 i 个行人对车辆的地图构建的矩阵, $\alpha_{mn}(m, n)$ 用来判断第 j 个车辆是否在构建的地图内。因此步骤A中LSTM层C的输出为:

$$[0042] \quad v_c^{i,t} = LSTM(v_c^{i,t-1}, C_t^i; W_3) \quad (13)$$

[0043] $v_c^{i,t}$ 是行人 i 在 t 时刻的LSTM层C的输出, W_3 是LSTM层C的权重。

[0044] C、建立人车交互的LSTM神经网络的输出

[0045] 将步骤B中LSTM层A、LSTM层B和LSTM层C的三个输出链接到一起输入到LSTM层D中, 输出预测的行人轨迹, 如下式所示:

$$[0046] \quad h_t^i = p_t^i \oplus s_t^i \oplus v_t^i \quad (14)$$

$$[0047] \quad r_c^{i,t} = LSTM(r_c^{i,t-1}, h_t^i; W_4) \quad (15)$$

$$[0048] \quad (\hat{x}_t^i, \hat{y}_t^i) = W_o h_t + b_o \quad (16)$$

[0049] 其中 \oplus 表示链接运算符, h_t^i 是LSTM层D的输入, r_t^i 是LSTM层D的输出, W_4 是LSTM层D的权重。 \hat{x}_t^i 、 \hat{y}_t^i 是最终预测到的行人轨迹坐标, W_o 、 b_o 分别是LSTM层D的权重和偏移量。

[0050] 本发明的效果和益处是:

[0051] 1、本发明选用行人的前进方向和速度作为输入,更直观的表现出邻居和车辆对当前行人运动的影响。

[0052] 2、本发明引入邻居行人和车辆的影响作为“社会信息”,与行人轨迹共同作为输入,运用分层编码的模式构建神经网络,能够解决行人受社会因素影响所发生轨迹改变的问题,提高预测精度。

[0053] 3、本发明针对于人-车交互的“社会信息”方面,提出方向注意力函数,用来区分不同方向的车辆对行人的影响,以提高“社会信息”的精度,进而提高行人轨迹预测精度。

附图说明

[0054] 本发明共有附图3张,其中:

[0055] 图1是网络基本结构图。

[0056] 图2是人-人栅格地图构建图。

[0057] 图3是人-车栅格地图构建图。

具体实施方式

[0058] 下面结合附图对本发明进行进一步地描述。以DUT环形路口数据集为例对本发明进行进一步的分析和说明。

[0059] A、按图1所示,构建人车交互的LSTM神经网络;

[0060] B、建立多层神经网络的输入;

[0061] B1、输入当前行人轨迹;

[0062] B2、输入人-人交互信息;

[0063] 将步骤B2中的扇形区域划分为如图2所示的 4×4 共16个网格,构成人-人栅格地图,人-人栅格地图内的数字即每个网格内的人数,因此可以将栅格地图写成 4×4 的矩阵的形式,不同网格之内的行人的数量构成 4×4 的矩阵中的元素;

[0064] B3、输入人-车交互信息;

[0065] 将步骤B3中圆形地图分为 2×2 大小的四个网格,构建人-车栅格地图;

[0066] C、建立人车交互的LSTM神经网络的输出;

[0067] D、评价预测结果;

[0068] 用训练神经网络的预测值与目标值的均方误差MSE来评价神经网络预测的精确度,MSE的值越小,预测的精确度越高。当MSE值小于设定的误差目标值,即完成对神经网络的训练,若MSE值大于设定的误差目标值,需重新调整神经网络参数,再进行训练。

[0069] 本发明不局限于本实施例,任何在本发明披露的技术范围内的等同构思或者改变,均列为本发明的保护范围。

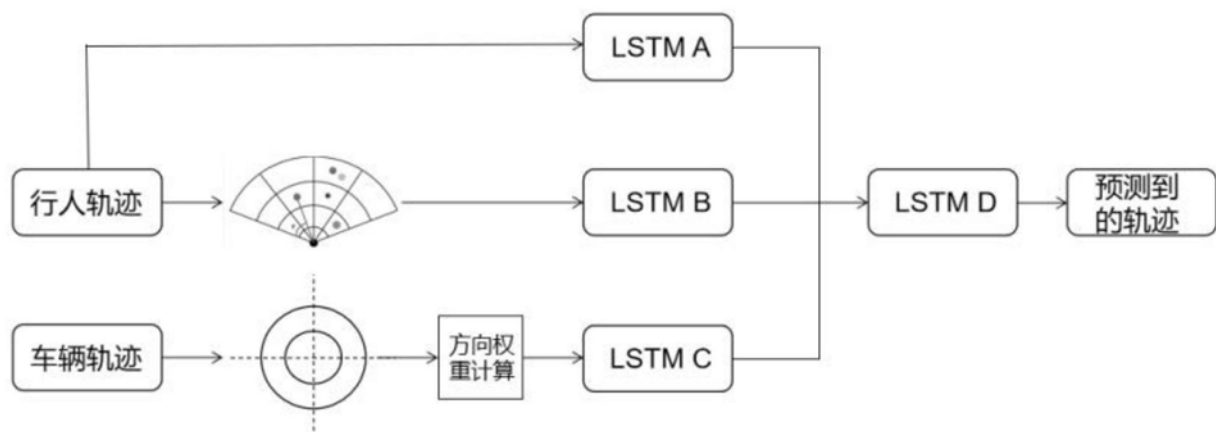


图1

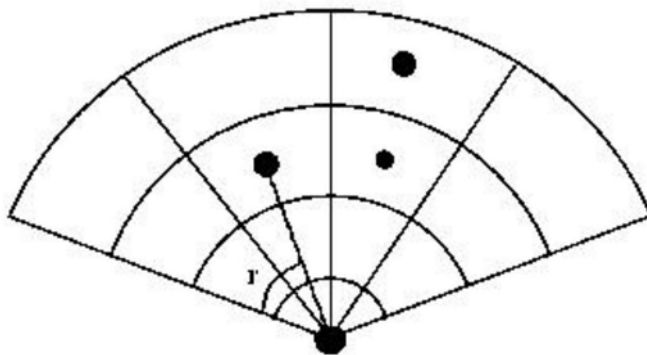


图2

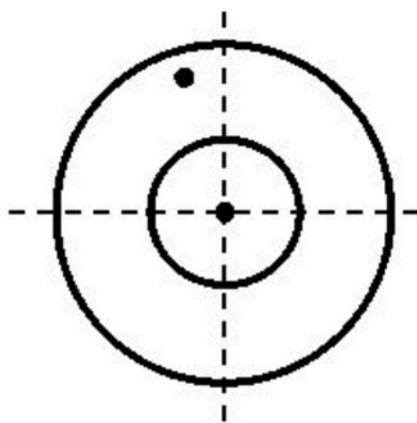


图3