



# (12) 发明专利

(10) 授权公告号 CN 112699785 B

(45) 授权公告日 2022.06.07

(21) 申请号 202011601643.3

CN 107220591 A, 2017.09.29

(22) 申请日 2020.12.29

CN 111368649 A, 2020.07.03

(65) 同一申请的已公布的文献号

CN 111914594 A, 2020.11.10

申请公布号 CN 112699785 A

CN 104732203 A, 2015.06.24

US 10061977 B1, 2018.08.28

(43) 申请公布日 2021.04.23

US 2016287166 A1, 2016.10.06

(73) 专利权人 中国民用航空飞行学院

张严浩. 基于结构化认知计算的群体行为分析.《中国博士学位论文全文数据库(信息科技辑)》.2018,

地址 618000 四川省德阳市广汉市南昌路四段46号

张婷. 基于PAD三维情感模型的情感语音研究.《中国优秀硕士学位论文全文数据库(信息科技辑)》.2018,

(72) 发明人 潘磊 王艾 赵欣 刘国春

高大鹏 袁小珂 严宏 马婷

朱建刚 严崇耀 卢志伟

饶元等. 基于语义分析的情感计算技术研究进展.《软件学报》.2018, 第29卷(第8期),

(74) 专利代理机构 西安铭泽知识产权代理事务所(普通合伙) 61223

Reda Elbarougy等. Continuous Audiovisual Emotion Recognition Using Feature Selection and LSTM.《Journal of Signal Processing》.2020, 第24卷(第6期),

专利代理师 王力文

Shizhe Chen等. Multimodal Multi-task Learning for Dimensional and Continuous Emotion Recognition.《AVEC'17: Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge》.2017,

(51) Int. Cl.

G06V 40/20 (2022.01)

G06V 40/16 (2022.01)

G06V 20/40 (2022.01)

G06V 10/764 (2022.01)

G06K 9/62 (2022.01)

审查员 刘利

(56) 对比文件

CN 107169426 A, 2017.09.15

权利要求书3页 说明书10页 附图6页

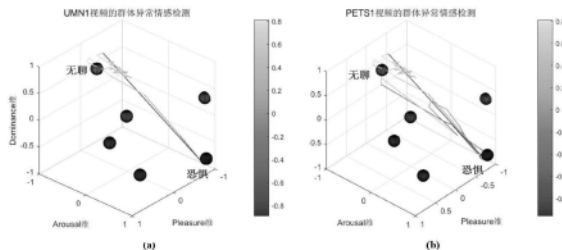
## (54) 发明名称

基于维度情感模型的群体情感识别与异常情感检测方法

运动视频, 本申请既能够准确表达群体情感的连续变化状态, 还可以实现对全局异常状态的有效识别。

## (57) 摘要

本发明公开了基于维度情感模型的群体情感识别与异常情感检测方法, 涉及智能情感识别技术领域, 基于认知心理学PAD三维情感模型, 通过数据搜集和人工标注, 创建群体情感的视频数据集, 揭示了六种典型情感在PAD空间中的位置关系; 创建基于群体行为的情感预测模型, 将群体运动特征映射为PAD空间中的三维坐标; 构建异常情感分类器, 当检测到愤怒和恐惧这两种异常情感, 则判定场景产生了异常状态。针对群体



CN 112699785 B

1. 基于维度情感模型的群体情感识别与异常情感检测方法, 其特征在于, 包括以下步骤:

S1: 建立基于群体情感的PAD三维情感模型: 包括愉悦度P、激活度A和优势度D三个维度, 每个维度的取值均处于-1到+1之间, 设置PAD情绪量表进行情绪维度的参考;

S2: 建立群体行为和群体情感数据集: 针对不同场景的视频数据, 基于认知心理学原理, 通过人工标注的策略, 获取标准视频数据集;

S3: 统计群体情感数据集: 根据标准视频数据集, 定义视频的情感类型, 被标记为同一类情感的视频, 将其PAD值归一化到[-1, 1]之间, 通过计算各个坐标的中心点, 确定该类情感在PAD空间中的取值;

S4: 评价群体情感数据集: 检验标注数据是否具有 consistency, 采用Normplot函数验证分析标注数据是否服从高斯分布, 若不服从则输出图像出现弯曲;

S5: 群体情感识别与异常情感检测: 从视频中提取群体运动特征, 群体运动中层语义的表达;

其中, 所述提取群体运动特征包括前景区域的提取、光流特征的提取、轨迹特征的提取和运动特征的图形化表达; 所述前景区域的提取采用改进的ViBE+算法, 经检测, 第t帧的前景区域表示为 $R^t$ ; 所述光流特征的提取采用Gunner Farneback的稠密光流场进行视觉表达, 对于第t帧图像, 像素点(x, y)在横向和纵向的光流偏移量分别为u和v; 所述轨迹特征的提取采用了iDT算法, 对视频像素点进行密集采集, 并通过光流判断跟踪点在下一帧的位置, 从而形成跟踪轨迹, 表示为 $T(p_1, p_2, \dots, p_L)$ , 其中 $L \leq 15$ ; 所述运动特征的图形化表达采用全局运动强度图、全局运动方向图和全局运动轨迹图三种图形化的特征表达形式;

全局运动轨迹图中的每条轨迹均以实线表示, 每条轨迹包括三个属性特征 $\langle T(p_1, p_2, \dots, p_L), L, g_i \rangle$ ; 其中,  $T(p_1, p_2, \dots, p_L)$ 表示构成轨迹的若干跟踪点 $p_i$ 的坐标位置, L表示轨迹的长度,  $g \in [0, 255]$ 表示轨迹中第i段的灰度值,  $g_i$ 表示如下:

$$g_i = \left\lceil \frac{255 \times (i+1)}{L+1} \right\rceil \quad (2)$$

其中,  $i \in [1, L-1]$ ;

所述步骤S5群体运动中层语义的表达采用灰度共生矩阵进行深入分析, 采用的统计量包括方差、对比度、二阶矩、熵、相关性和倒数差分矩;

所述方差用于反映图像的灰度变化程度, 当方差越大时, 图像的灰度变化越大, 方差的计算公式如下:

$$Var = \sum_i \sum_j p(i, j) \times (i - Meam)^2 \quad (3)$$

其中,  $Meam = \sum_i \sum_j p(i, j) \times i$ ;

所述对比度用于度量矩阵的值分布和图像中局部的变化量, 用于反映图像的清晰度和纹理的深浅, 对比度的计算公式如下:

$$Con = \sum_i \sum_j (i, j)^2 p(i, j) \quad (4);$$

所述二阶矩用于度量图像纹理的灰度变化稳定程度, 反映图像灰度分布均匀程度和纹

理粗细度,二阶矩值较大则表明处于较均一和规则变化的纹理模式,二阶矩的计算公式如下:

$$ASM = \sum_i \sum_j p(i, j)^2 \quad (5);$$

所述熵用于度量图像包含信息量的随机性,反映图像灰度分布的复杂程度,熵的计算公式如下:

$$Ent = -\sum_i \sum_j p(i, j) \ln p(i, j) \quad (6);$$

所述相关性用于度量空间灰度共生矩阵元素在行或列方向上的相似程度,反映图像纹理的一致性,相关性的计算公式如下:

$$Corr = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j) p(i, j)}{\sigma_i \sigma_j} \quad (7);$$

所述倒数差分矩用于反映图像纹理的同质性,度量图像纹理局部变化的多少,其值大则说明图像纹理的不同区域间缺少变化,局部均匀,倒数差分矩的计算公式如下:

$$IDM = \sum_i \sum_j \frac{p(i, j)}{1 + (i - j)^2} \quad (8);$$

S6:群体情感特征的提取、回归:采用支持向量回归机SVR,在训练数据集的支持下,通过寻找最优超平面,在约束其结构化风险最小化的基础上,获得回归函数;所述回归函数如下所示:

$$\min_{w, b, \xi^v, \xi^{\wedge}} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N (\xi_i + \widehat{\xi}_i) \quad (9)$$

$$s.t. \begin{cases} y_i - \omega \times \varphi(x_i) - b \leq \widehat{\xi}_i + \varepsilon \\ \omega \times \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i \\ \xi_i \geq 0, \widehat{\xi}_i \geq 0, i = 1, 2, \dots, N \end{cases} \quad (10)$$

其中, $\omega$ 为权重向量, $C$ 为平衡系数, $\widehat{\xi}_i$ 、 $\xi_i$ 为松弛变量, $\varphi(x)$ 为将数据映射到高维空间的非线性变换, $b$ 为偏置项, $\varepsilon$ 为敏感度;

引入拉格朗日乘子,式(10)转换为:

$$\min \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\widehat{\alpha}_i - \alpha_i) (\widehat{\alpha}_j - \alpha_j) k(x, x_i) + \sum_{i=1}^N (\varepsilon - y_i) \cdot \widehat{\alpha}_j + (\varepsilon + y_i) \cdot \widehat{\alpha}_i \quad (11)$$

其中, $0 \leq \alpha_i \leq C, 0 \leq \widehat{\alpha}_i \leq C, 0 \leq a_i \leq C$ ;

最终求得的回归函数为:

$$f(x) = \sum_{i=1}^N (a_i - \widehat{a}_i) k(x, x_i) + b \quad (12)$$

其中, $k(x, x_i)$ 为核函数;

采用径向基核函数RBF,表达式如下:

$$k(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2) \quad (13)$$

通过训练后得到回归模型,实现维度情感预测,对每段视频在PAD空间中的连续值进行预测,当群体情感随时间发生变化,则表现为连续的三维轨迹,以呈现渐变的情感过程;

S7:异常情感状态的检测:将每个标记样本的PAD值作为输入,通过支持向量机SVM进行训练。

2.如权利要求1所述的基于维度情感模型的群体情感识别与异常情感检测方法,其特征在于,所述步骤S2中根据人工标注的策略设计情感标注系统,该系统通过人物模型的面部表情表示P维取值,以心脏的震动程度表示A维取值,以小人的大小来表示D维取值。

3.如权利要求1所述的基于维度情感模型的群体情感识别与异常情感检测方法,其特征在于,所述步骤S4判断一致性的方法如下:计算变异系数,统计和评价PAD数据的样本均值 $\mu$ 、样本标准差 $\sigma$ 以及变异系数CV三个指标,其中,变异系数的定义为:

$$CV = \left| \frac{\sigma}{\mu} \right| \times 100\% \quad (1)$$

若变异系数小,则说明验证标注数据一致性低;反之,则说明验证标注数据一致性高。

4.如权利要求1所述的基于维度情感模型的群体情感识别与异常情感检测方法,其特征在于,所述步骤S7的异常情感状态的检测获得SVM超平面的二次方程,其表达式为:

$$\min_{w, \xi, p} \frac{1}{2} \|w\|^2 + \frac{1}{vN} \sum_{i=1}^N \xi_i - p \quad (14)$$

其中, $s \cdot t \cdot$ ,  $w^T \Phi(x_i) \geq \rho - \xi_i$ ,  $\xi_i \geq 0$ ,  $x_i$ 表示 $i = \{1, 2, \dots, N\}$ 的训练集数据 $w^T \Phi(x_i) - \rho = 0$ 最大决策超平面; $\xi_i$ 为惩罚异常点的松弛变量; $v \in (0, 1]$ 为百分比估计; $\Phi(\cdot)$ 为训练数据映射到高维特征空间的非线性方程;此外,径向基核函数定义为 $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ ,在特征空间内进行点乘运算,采用高斯核函数,决策函数定义为:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i k(x_i, x_j) - b \right) \quad (15)。$$

## 基于维度情感模型的群体情感识别与异常情感检测方法

### 技术领域

[0001] 本发明涉及智能情感识别技术领域,特别涉及基于维度情感模型的群体情感识别与异常情感检测方法。

### 背景技术

[0002] 近年来,随着人工智能与深度学习、心理科学及认知科学的不断发展,通过计算机来识别、理解、表达和交流人的情感,使计算机具备更全面、更高层次的智能化程度,越来越受到学术界的广泛关注和深入探索。对于智能视频监控技术而言,通过采集场景中人群的语言交流、面部表情、肢体运动,理解和体会他们的喜怒哀乐,分析其情感状态和内心意图,并推断其下一步的行为企图,使计算机随之做出相应的反馈,从而真正具备情感层面的交流能力。作为未来智能监控技术的重要发展方向之一,基于视觉的情感分析和情感识别方法,具有十分重要的学术研究价值。

[0003] 当然,人类表达情感的信息媒介有多种渠道,包括文字、语言、面部表情、肢体行为等。虽然语言和面部表情都可以丰富地表达人类情感,但是语音信号在嘈杂的公共场所中很难清晰采集到。同时,考虑到密集场景的高度拥挤和动态变化,现有的视频分析技术很难从拥挤的人群中准确定位每个人的脸部,并精确提取面部表情。因此,基于人脸跟踪和面部特征提取的情感分析,在密集场景中很难达到理想效果。所以,比较可行的途径,是通过分析监控视频中人群的肢体行为来识别和评价群体情感状态。

[0004] 值得注意的是,目前学术界对于肢体运动的情感分析,往往将单独的个人作为研究对象,侧重于挖掘和识别个体姿态特征及其情感表达。然而,区别于个体运动,群体行为在主观因素、环境因素、社会因素和心理因素的共同作用下,具有其独特的内在结构和丰富的外在形态。一方面,个体之间通过信息交流和相互协作,使群体呈现出一定的趋同性和整体性;但另一方面,个体运动具有一定的自主性和随意性,群体又表现出一定的无序性和非结构化特点。从社会心理学的角度分析,在密集人群场景中,个体心理受周围环境影响,会丧失一定的独立性,对同伴形成某种程度的依赖,导致其情绪状态逐渐与群体趋于一致,形成集体的、从众的心理状态。因此,考虑到密集场景的特殊性以及群体心理的独特性,有必要探索专门的方法和策略来分析群体情感状态。

[0005] 基于群体行为的情感识别方法,目前主要分为两种类型:基于离散模型模型的识别方法和基A-V二维情感模型的识别方法。然而这两类方法都存在一些不足之处。第一,不同于简单的一段语音或者一幅图像,监控视频所呈现的内容是非常丰富的。这其中既有活跃的群体运动,又有复杂的群体情感,还有一定的情节变化。所以,离散情感模型只能识别一些形态单一、辨识度高的典型场景,其所涵盖的几种特定情感类型对密集人群来说,是比较有限和欠缺的。此外,群体情感具有许多微妙的特征,表现为多种情绪的综合。而且情感会随时间发生连续的变化。群体情感这些特点,都是离散模型无法有效表达的。第二,A-V二维情感模型主要从Arousal和Valence两个维度加以衡量。其中,Arousal反映了情感状态的强度,Valence体现了情感状态的类型。但是相比于三维情感模型,两个维度的描述形式

还是稍显简单,例如文献采用A-V二维情感模型,仅仅区分了四种情感类别。对于复杂的群体情感来说,这显然是不够的。第三,A-V情感模型无法区分某些情感(如愤怒和恐惧都属于Arousal较高的情感),但是PAD三维情感模型可以有效区分(愤怒属于优势度较高的情感,而恐惧属于优势度较低的情感)。

[0006] 针对上述问题,本申请提供了一种基于维度情感模型的群体情感识别与异常情感检测方法,以PAD维度模型为基础,将群体情感表示为情感空间中的三维坐标点,以实现复杂情感的准确表达。

### 发明内容

[0007] 本发明的目的在于提供基于维度情感模型的群体情感识别与异常情感检测方法,以PAD维度模型为基础,将群体情感表示为情感空间中的三维坐标点,以实现复杂情感的准确表达。

[0008] 本发明提供了基于维度情感模型的群体情感识别与异常情感检测方法,包括以下步骤:

[0009] S1:建立基于群体情感的PAD三维情感模型:包括愉悦度P、激活度A和优势度D三个维度,每个维度的取值均处于-1到+1之间,设置PAD情绪量表进行情绪维度的参考;

[0010] S2:建立群体行为和群体情感数据集:针对不同场景的视频数据,基于认知心理学原理,通过人工标注的策略,获取标准视频数据集;

[0011] S3:统计群体情感数据集:根据标准视频数据集,定义视频的情感类型,被标记为同一类情感的视频,将其PAD值归一化到[-1,1]之间,通过计算各个坐标的中心点,确定该类情感在PAD空间中的取值;

[0012] S4:评价群体情感数据集:检验标注数据是否具有 consistency,采用Matlab工具中的Normplot函数验证分析标注数据是否服从高斯分布,若不服从则输出图像出现弯曲;

[0013] S5:群体情感识别与异常情感检测:从视频中提取群体运动特征,群体运动中层语义的表达;

[0014] S6:群体情感特征的提取、回归:采用支持向量回归机SVR,在训练数据集的支持下,通过寻找最优超平面,在约束其结构化风险最小化的基础上,获得回归函数;

[0015] S7:异常情感状态的检测:将每个标记样本的PAD值作为输入,通过支持向量机SVM进行训练。

[0016] 进一步地,所述步骤S2中根据人工标注的策略设计情感标注系统,该系统通过人物模型的面部表情表示P维取值,以心脏的震动程度表示A维取值,以小人的大小来表示D维取值。

[0017] 进一步地,所述步骤S4判断一致性的方法如下:计算变异系数,统计和评价PAD数据的样本均值 $\mu$ 、样本标准差 $\sigma$ 以及变异系数CV三个指标,其中,变异系数的定义为:

$$[0018] \quad CV = \left| \frac{\sigma}{\mu} \right| \times 100\% \quad (1)$$

[0019] 若变异系数小,则说明验证标注数据一致性低;反之,则说明验证标注数据一致性高。

[0020] 进一步地,所述步骤S5提取群体运动特征包括前景区域的提取、光流特征的提取、

轨迹特征的提取和运动特征的图形化表达;所述前景区域的提取采用改进的ViBE+算法,经检测,第t帧的前景区域表示为 $R^t$ ;所述光流特征的提取采用Gunner Farneback的稠密光流场进行视觉表达,对于第t帧图像,像素点 $(x, y)$ 在横向和纵向的光流偏移量分别为 $u$ 和 $v$ ;所述轨迹特征的提取采用了iDT算法,对视频像素点进行密集采集,并通过光流判断跟踪点在下一帧的位置,从而形成跟踪轨迹,表示为 $T(p_1, p_2 \dots p_L)$ ,其中 $L \leq 15$ ;所述运动特征的图形化表达采用全局运动强度图、全局运动方向图和全局运动轨迹图三种图形化的特征表达形式。

[0021] 进一步地,全局运动轨迹图中的每条轨迹均以实线表示,每条轨迹包括三个属性特征 $\langle T(p_1, p_2 \dots p_L), L, g_i \rangle$ ;其中, $T(p_1, p_2 \dots p_L)$ 表示构成轨迹的若干跟踪点 $p_i$ 的坐标位置, $L$ 表示轨迹的长度, $g \in [0, 255]$ 表示轨迹中第 $i$ 段的灰度值, $g_i$ 表示如下:

$$[0022] \quad g_i = \left\lceil \frac{255 \times (i+1)}{L+1} \right\rceil \quad (2)$$

[0023] 其中, $i \in [1, L-1]$ 。

[0024] 进一步地,所述步骤S5群体运动中层语义的表达采用灰度共生矩阵进行深入分析,采用的统计量包括方差、对比度、二阶矩、熵、相关性和倒数差分矩;

[0025] 所述方差用于反映图像的灰度变化程度,当方差越大时,图像的灰度变化越大,方差的计算公式如下:

$$[0026] \quad Var = \sum_i \sum_j p(i, j) \times (i - Meam)^2 \quad (3)$$

[0027] 其中, $Meam = \sum_i \sum_j p(i, j) \times i$ ;

[0028] 所述对比度用于度量矩阵的值分布和图像中局部的变化量,用于反映图像的清晰度和纹理的深浅,对比度的计算公式如下:

$$[0029] \quad Con = \sum_i \sum_j (i, j)^2 p(i, j) \quad (4);$$

[0030] 所述二阶矩用于度量图像纹理的灰度变化稳定程度,反映图像灰度分布均匀程度和纹理粗细度,二阶矩值较大则表明处于较均一和规则变化的纹理模式,二阶矩的计算公式如下:

$$[0031] \quad ASM = \sum_i \sum_j p(i, j)^2 \quad (5);$$

[0032] 所述熵用于度量图像包含信息量的随机性,反映图像灰度分布的复杂程度,熵的计算公式如下:

$$[0033] \quad Ent = -\sum_i \sum_j p(i, j) \ln p(i, j) \quad (6);$$

[0034] 所述相关性用于度量空间灰度共生矩阵元素在行或列方向上的相似程度,反映图像纹理的一致性,相关性的计算公式如下:

$$[0035] \quad Corr = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} p(i, j) \quad (7);$$

[0036] 所述倒数差分矩用于反映图像纹理的同质性,度量图像纹理局部变化的多少,其值大则说明图像纹理的不同区域间缺少变化,局部均匀,倒数差分矩的计算公式如下:

$$[0037] \quad IDM = \sum_i \sum_j \frac{p(i,j)}{1+(i-j)^2} \quad (8)。$$

[0038] 进一步地,所述步骤S6回归函数如下所示:

$$[0039] \quad \min_{w,b,\xi^v,\xi^{\wedge}} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N (\xi_i + \widehat{\xi}_i) \quad (9)$$

$$[0040] \quad s \cdot t \cdot \begin{cases} y_i - \omega \times \varphi(x_i) - b \leq \widehat{\xi}_i + \varepsilon \\ \omega \times \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i \\ \xi_i \geq 0, \widehat{\xi}_i \geq 0 (i=1,2,\dots,N) \end{cases} \quad (10)$$

[0041] 其中, $\omega$ 为权重向量, $C$ 为平衡系数, $\widehat{\xi}_i$ 、 $\xi_i$ 为松弛变量, $\varphi(x)$ 为将数据映射到高维空间的非线性变换, $b$ 为偏置项, $\varepsilon$ 为敏感度;

[0042] 引入拉格朗日乘子,式(10)转换为:

$$[0043] \quad \min \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\widehat{\alpha}_i - \alpha_i)(\widehat{\alpha}_j - \alpha_j) k(x, x_j) + \sum_{i=1}^N (\varepsilon - y_i) \cdot \widehat{\alpha}_i + (\varepsilon + y_i) \cdot \alpha_i \quad (11)$$

$$[0044] \quad \text{其中, } s \cdot t \cdot \sum_{i=1}^N (\widehat{\alpha}_i - \alpha_i) = 0, \quad 0 \leq \widehat{\alpha}_i \leq C, \quad 0 \leq \alpha_i \leq C;$$

[0045] 最终求得的回归函数为:

$$[0046] \quad f(x) = \sum_{i=1}^N (a_i - \widehat{a}_i) k(x, x_i) + b \quad (12)$$

[0047] 其中, $k(x, x_i)$ 为核函数;

[0048] 采用径向基核函数RBF,表达式如下:

$$[0049] \quad k(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2) \quad (13)$$

[0050] 通过训练后得到回归模型,实现维度情感预测,对每段视频在PAD空间中的连续值进行预测,当群体情感随时间发生变化,则表现为连续的三维轨迹,以呈现渐变的情感过程。

[0051] 进一步地,所述步骤S7的异常情感状态的检测获得SVM超平面的二次方程,其表达式为:

$$[0052] \quad \min_{w,\xi,p} \frac{1}{2} \|w\|^2 + \frac{1}{vN} \sum_{i=1}^N \xi_i - p \quad (14)$$

[0053] 其中, $s \cdot t \cdot w^T \Phi(x_i) \geq p - \xi_i, \xi_i \geq 0, x_i$ 表示 $i = \{1, 2, \dots, N\}$ 的训练集数据 $w^T \Phi(x_i) - p = 0$ 最大决策超平面; $\xi_i$ 为惩罚异常点的松弛变量; $v \in (0, 1]$ 为百分比估计; $\Phi(\cdot)$ 为训练数据映射到高维特征空间的非线性方程;此外,核函数定义为 $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ ,在特征空间内进行点乘运算,采用高斯核函数,决策函数定义为:

$$[0054] \quad f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i k(x_i, x_j) - b \right) \quad (15)。$$



[0055] 与现有技术相比,本发明具有如下显著优点:

[0056] 本申请提供了一种基于维度情感模型的群体情感识别与异常情感检测方法,第一,首次将三维情感模型应用于密集人群场景下的群体情感识别,以PAD维度模型为基础,将群体情感表示为情感空间中的三维坐标点,以实现复杂情感的准确表达。

[0057] 第二,首次创建了面向群体行为的维度情感数据集,通过人工标注和统计分析的方法,揭示了各类情感在三维情感空间中的坐标与联系,从而为后续的情感分析奠定数据基础。

[0058] 第三,提出了一系列从群体运动中提取情感特征的方法。在维度情感的有关定义下,通过支持向量回归,构建了从运动到情感的抽象过程和映射方法。

[0059] 第四,将惊恐和愤怒两种情感定义为异常情感。通过识别这两种情感,可以判断场景产生了异常状态。这就为场景的智能化检测,从情感识别的角度,开辟了一种新颖的解决途径。

### 附图说明

[0060] 图1为本发明实施例提供的基于UMN和PET2009数据集的群体异常情感检测图;

[0061] 图2为本发明实施例提供的视频片段分别在PAD维的数据分析图;

[0062] 图3为本发明实施例提供的群体情感识别和异常状态检测流程图;

[0063] 图4为本发明实施例提供的群体运动特征的提取和中层语义表达图;

[0064] 图5为本发明实施例提供的GMIC、GMOC和GMTC效果图;

[0065] 图6为本发明实施例提供的异常情感状态的检测流程图;

[0066] 图7为本发明实施例提供的六种情感类型在PAD维度空间中的坐标图。

### 具体实施方式

[0067] 下面结合本发明中的附图,对本发明实施例的技术方案进行清楚、完整的描述,显然,所描述的实施例是本发明的一部分实施例,而不是全部实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都应属于本发明保护的范围。

[0068] 心理学理论认为,人的内心情感与外在行为之间存在着很强的关联性,下意识的肢体动作能够流露出人内心深处的情感和意图。因此,在心理学模型和社会情感原则的基础上,通过描述群体行为的情感属性来识别情感状态是可行的。从宏观的角度分析,群体的不同形态和运动模式,往往从整体上反映了许多典型的情感状态。

[0069] 根据现有的文献资料分析,从心理学模型的角度来划分,基于群体行为的情感识别方法,目前主要分为两种类型:基于离散模型模型的识别方法和基于A-V二维情感模型的识别方法。然而这两类方法都存在一些不足之处。

[0070] 第一,不同于简单的一段语音或者一幅图像,监控视频所呈现的内容是非常丰富的。这其中既有活跃的群体运动,又有复杂的群体情感,还有一定的情节变化。所以,离散情感模型只能识别一些形态单一、辨识度高的典型场景,其所涵盖的几种特定情感类型对密集人群来说,是比较有限和欠缺的。此外,群体情感具有许多微妙的特征,表现为多种情绪的综合。而且情感会随时间发生连续的变化。群体情感这些特点,都是离散模型无法有效

表达的。

[0071] 第二,A-V二维情感模型主要从Arousal和Valence两个维度加以衡量。其中,Arousal反映了情感状态的强度,Valence体现了情感状态的类型。但是相比于三维情感模型,两个维度的描述形式还是稍显简单。

[0072] 第三,A-V情感模型无法区分某些情感(如愤怒和恐惧都属于Arousal较高的情感),但是PAD三维情感模型可以有效区分(愤怒属于优势度较高的情感,而恐惧属于优势度较低的情感)。

[0073] 本申请提出的方法从群体情感识别的角度,基于心理学PAD三维情感模型,通过数据搜集和人工标注,创建了群体情感的视频数据集,揭示了六种典型情感在PAD空间中的位置关系。构建了基于群体行为的情感预测模型,将群体运动特征映射为PAD空间中的三维坐标。包括:运动特征提取和中层语义表达,群体情感特征提取和融合,以及情感状态回归。构建了异常情感分类器,当检测到愤怒和恐惧这两种异常情感,则判定场景产生了异常状态。本申请提供的方法既能够准确表达群体情感的连续变化状态,还可以实现对全局异常状态的有效识别。

[0074] 参照图1-7,本发明提供了基于维度情感模型的群体情感识别与异常情感检测方法,包括以下步骤:

[0075] S1:建立基于群体情感的PAD三维情感模型:包括愉悦度P、激活度A和优势度D三个维度,每个维度的取值均处于-1到+1之间,设置PAD情绪量表进行情绪维度的参考;

[0076] S2:建立群体行为和群体情感数据集:针对不同场景的视频数据,基于认知心理学原理,通过人工标注的策略,获取标准视频数据集;

[0077] S3:统计群体情感数据集:根据标准视频数据集,定义视频的情感类型,被标记为同一类情感的视频,将其PAD值归一化到[-1,1]之间,通过计算各个坐标的中心点,确定该类情感在PAD空间中的取值;

[0078] S4:评价群体情感数据集:检验标注数据是否具有有一致性,采用Matlab工具中的Normplot函数验证分析标注数据是否服从高斯分布,若不服从则输出图像出现弯曲;

[0079] S5:群体情感识别与异常情感检测:从视频中提取群体运动特征,群体运动中层语义的表达;

[0080] S6:群体情感特征的提取、回归:采用支持向量回归机SVR,在训练数据集的支持下,通过寻找最优超平面,在约束其结构化风险最小化的基础上,获得回归函数;

[0081] S7:异常情感状态的检测:将每个标记样本的PAD值作为输入,通过支持向量机SVM进行训练。

[0082] 其中,步骤S1中P表示个体情绪状态的正负特性,包括情感的积极或消极两个对立的状态,如喜欢或不喜欢,满意或不满意,愉快或不高兴。愉悦度取值为正数,则表示正面情感,反之则表示负面情感。一般来说,群体运动的不同类型可以表达密集人群的正面或负面情感。例如,缓缓步行、驻足交谈,都体现了正面情感;而斗殴冲突、四散跑开则是典型的负面情感。

[0083] A表示个体的神经生理激活水平、警觉性,与情感状态联系的机体能量的激活程度有关,也就是情绪的强度特征,包括低唤醒状态(如安静)和高唤醒状态(如惊奇)两个状态。对于密集人群来说,运动变化的剧烈程度体现了激活度水平的高低。例如,当人群自由前行

突然变化为四散逃亡,在一般情况下,这表明人群受某种外界因素的刺激和影响所导致,人群的激活度从低唤醒状态转变为高唤醒状态。

[0084] D表示个体对情景和他人的控制状态,主要指个体对情感状态的主观控制程度,用以区分情感状态是由个体主观发出的还是受客观环境影响产生的。对于密集人群来说,个体运动的差异性和同质性代表了优势度的高低。当个体运动体现出了一定自主性、随意性和无序性,例如在广场、街道上悠闲步行的人群,其个体行为主要遵从于自身主观意识,则优势度较高。当个体运动体现出一定的从众、趋同的特点,例如在疏散逃亡的人群,所有人都在往某一个方向奔跑,则个体运动受制于群体运动模式,群体运动模式在宏观上趋于一致,则优势度较低。

[0085] 实施例1

[0086] 所述步骤S2中根据人工标注的策略设计情感标注系统,该系统通过人物模型的面部表情表示P维取值,以心脏的震动程度表示A维取值,以小人的大小来表示D维取值。

[0087] 情感数据集的构造方法主要有两种:演绎方式和摘引方式。演绎方式是由

[0088] 表演者(最好具有专业表演素养)通过肢体动作来模拟某种典型情感类型(喜悦、惊恐、悲伤)。这种方式的情感对比度鲜明、表现力较强,但这种演绎形式与真实的情感尚存在一定的差距,且对表演者的表演素养要求较高,不具备普遍性。而摘引的方式,是从现实场景的视频片段中,采用人工标注的方法,对群体行为的情感状态及其所示,各项指标评价打分。由此获取到的情感是人的自然流露,更接近于现实生活,但后期标注的工作量较大。

[0089] 目前学术界尚缺乏群体行为的情感数据库。有部分学者在研究工作中,进行了类似的工作,提出了一些有标定的数据集,但尚未公开,我们无法验证其有效性。现有技术的主要用途是为了提升群体行为的检测效率,而非情感分析。是针对个体姿态和个体行为的数据集,鉴于此,本申请开展了情感标注的实验,从而建立起群体行为和群体情感的数据集,针对不同的场景,摘引若干真实的视频数据中,通过第三方人工标注的策略,获取标准视频数据集。

[0090] 情感实验所构建的情感数据集来源于UMN数据集、PETS 2009数据集、UCF数据集、SCU数据集、UCF Crowd BEHAVE数据集、Web

[0091] Dataset:Abnormal/Normal Crowds数据集、Violent-flows数据集Rodriguez's数集中关于密集人群场景的视频,总计50个视频,以15帧为单位,截取了若干的视频片段,总计200个视频片段,邀请31位志愿者(17位男性、14位女性,年龄在19-35岁之间),分别对每个视频片段进行标注,包括两方面的工作:

[0092] (1) 志愿者要对每段视频的P、A、D三个维度值进行打分。每个维度的分值由低到高为{1,2,3,4,5}五个单选项。

[0093] (2) 志愿者要确定每段视频所表现的情感类型,包括{兴奋、愤怒、恐惧、平和、无聊、中性、以上都不是}七个单选项。

[0094] 实施例2

[0095] 所述步骤S4判断一致性的方法如下:计算变异系数,统计和评价PAD数据的样本均值 $\mu$ 、样本标准差 $\sigma$ 以及变异系数CV三个指标,其中,变异系数的定义为:

$$[0096] \quad CV = \left| \frac{\sigma}{\mu} \right| \times 100\% \quad (1)$$

[0097] 若变异系数小,则说明验证标注数据一致性低;反之,则说明验证标注数据一致性高。

[0098] 就不同的视频片段的PAD数据而言,统计其在同一维度下的标记值。如果变异系数较大,则反映其单位均值上的离散程度较大,说明志愿者对该组打分的一致性和确定性较低;反之,则说明志愿者对该组打分的一致性和确定性较高。通常来说,对于一致性较低的视频,如果变异系数大于20%,则要考虑该数据可能不正常,说明志愿者的分歧较大,可考虑将其从数据集中剔除,以保证数据的可信度。

[0099] 在此分别以视频片段为例,前一段视频表示的是人群四散奔逃,后一段视频表示的是人群在激烈打斗,关于这两段视频的PAD统计数据,从统计结果中可以看到,标注数据的变异系数CV都集中在[0,20%]区间,可以认为志愿者的打分都比较集中,分歧较小,因此这两组视频的PAD数据是可信的。

[0100] 实施例3

[0101] 所述步骤S5提取群体运动特征包括前景区域的提取、光流特征的提取、轨迹特征的提取和运动特征的图形化表达;所述前景区域的提取采用改进的ViBE+算法,经检测,第t帧的前景区域表示为 $R^t$ ;所述光流特征的提取采用Gunner Farnback的稠密光流场进行视觉表达,对于第t帧图像,像素点(x,y)在横向和纵向的光流偏移量分别为u和v;所述轨迹特征的提取采用了iDT算法,对视频像素点进行密集采集,并通过光流判断跟踪点在下一帧的位置,从而形成跟踪轨迹,表示为 $T(p_1, p_2 \cdots p_L)$ ,其中 $L \leq 15$ ;所述运动特征的图形化表达采用全局运动强度图、全局运动方向图和全局运动轨迹图三种图形化的特征表达形式。

[0102] 全局运动轨迹图中的每条轨迹均以实线表示,每条轨迹包括三个属性特征 $\langle T(p_1, p_2 \cdots p_L), L, g_i \rangle$ ;其中, $T(p_1, p_2 \cdots p_L)$ 表示构成轨迹的若干跟踪点 $p_i$ 的坐标位置,L表示轨迹的长度, $g_i \in [0, 255]$ 表示轨迹中第i段的灰度值, $g_i$ 表示如下:

$$[0103] \quad g_i = \left\lfloor \frac{255 \times (i+1)}{L+1} \right\rfloor \quad (2)$$

[0104] 其中, $i \in [1, L-1]$ 。

[0105] 所述步骤S5群体运动中层语义的表达采用灰度共生矩阵进行深入分析,采用的统计量包括方差、对比度、二阶矩、熵、相关性和倒数差分矩;

[0106] 所述方差用于反映图像的灰度变化程度,当方差越大时,图像的灰度变化越大,方差的计算公式如下:

$$[0107] \quad Var = \sum_i \sum_j p(i, j) \times (i - Meam)^2 \quad (3)$$

[0108] 其中, $Meam = \sum_i \sum_j p(i, j) \times i$ ;

[0109] 所述对比度用于度量矩阵的值分布和图像中局部的变化量,用于反映图像的清晰度和纹理的深浅,对比度的计算公式如下:

$$[0110] \quad Con = \sum_i \sum_j (i, j)^2 p(i, j) \quad (4);$$

[0111] 所述二阶矩用于度量图像纹理的灰度变化稳定程度,反映图像灰度分布均匀程度和纹理粗细度,二阶矩值较大则表明处于较均一和规则变化的纹理模式,二阶矩的计算公

式如下：

$$[0112] \quad ASM = \sum_i \sum_j p(i, j)^2 \quad (5) ;$$

[0113] 所述熵用于度量图像包含信息量的随机性,反映图像灰度分布的复杂程度,熵的计算公式如下：

$$[0114] \quad Ent = -\sum_i \sum_j p(i, j) \ln p(i, j) \quad (6) ;$$

[0115] 所述相关性用于度量空间灰度共生矩阵元素在行或列方向上的相似程度,反映图像纹理的一致性,相关性的计算公式如下：

$$[0116] \quad Corr = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} p(i, j) \quad (7) ;$$

[0117] 所述倒数差分矩用于反映图像纹理的同质性,度量图像纹理局部变化的多少,其值大则说明图像纹理的不同区域间缺少变化,局部均匀,倒数差分矩的计算公式如下：

$$[0118] \quad IDM = \sum_i \sum_j \frac{p(i, j)}{1 + (i - j)^2} \quad (8)。$$

[0119] 实施例4

[0120] 所述步骤S6回归函数如下所示：

$$[0121] \quad \min_{w, b, \xi^r, \xi^s} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N (\xi_i + \bar{\xi}_i) \quad (9)$$

$$[0122] \quad s \cdot t \cdot \begin{cases} y_i - \omega \times \varphi(x_i) - b \leq \bar{\xi}_i + \varepsilon \\ \omega \times \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i \\ \xi_i \geq 0, \bar{\xi}_i \geq 0 (i = 1, 2, \dots, N) \end{cases} \quad (10)$$

[0123] 其中,  $\omega$  为权重向量,  $C$  为平衡系数,  $\bar{\xi}_i$ 、 $\xi_i$  为松弛变量,  $\varphi(x)$  为将数据映射到高维空间的非线性变换,  $b$  为偏置项,  $\varepsilon$  为敏感度；

[0124] 引入拉格朗日乘子,式(10)转换为：

$$[0125] \quad \min \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\bar{\alpha}_i - \alpha_i)(\bar{\alpha}_j - \alpha_j) k(x, x_i) + \sum_{i=1}^N (\varepsilon - y_i) \cdot \bar{\alpha}_i + (\varepsilon + y_i) \cdot \bar{\alpha}_i \quad (11)$$

[0126] 其中,  $s \cdot t \cdot \sum_{i=1}^N (\bar{\alpha}_i - \alpha_i) = 0$ ,  $0 \leq \bar{\alpha}_i \leq C$ ,  $0 \leq \alpha_i \leq C$  ;

[0127] 最终求得的回归函数为：

$$[0128] \quad f(x) = \sum_{i=1}^N (a_i - \bar{a}_i) k(x, x_i) + b \quad (12)$$

[0129] 其中,  $k(x, x_i)$  为核函数；

[0130] 采用径向基核函数RBF,表达式如下：

$$[0131] \quad k(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2) \quad (13)$$

[0132] 通过训练后得到回归模型,实现维度情感预测,对每段视频在PAD空间中的连续值进行预测,当群体情感随时间发生变化,则表现为连续的三维轨迹,以呈现渐变的情感过

程。

[0133] 实施例5

[0134] 所述步骤S7的异常情感状态的检测获得SVM超平面的二次方程,其表达式为:

$$[0135] \quad \min_{w, \xi, \rho} \frac{1}{2} \|w\|^2 + \frac{1}{\nu N} \sum_{i=1}^N \xi_i - \rho \quad (14)$$

[0136] 其中,  $s \cdot t \cdot$ ,  $w^T \Phi(x_i) \geq \rho - \xi_i$ ,  $\xi_i \geq 0$ ,  $x_i$  表示  $i = \{1, 2, \dots, N\}$  的训练集数据  $w^T \Phi(x_i) - \rho = 0$  最大决策超平面;  $\xi_i$  为惩罚异常点的松弛变量;  $\nu \in (0, 1]$  为百分比估计;  $\Phi(\cdot)$  为训练数据映射到高维特征空间的非线性方程; 此外, 核函数定义为  $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ , 在特征空间内进行点乘运算, 采用高斯核函数, 决策函数定义为:

$$[0137] \quad f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i k(x_i, x_j) - b \right) \quad (15)。$$

[0138] 对于群体异常情感的检测结果, 六种情感类型在PAD空间中的坐标被确定。曲线的渐变色由浅入深, 表示了帧序列的先后顺序。由于视频中的群体情感状态发生了连续的变化, 这表现为表示群体情感状态的连续变化。视频中群体情感随时间所发生的变化过程, 在图中表现为连续的情感轨迹。可以看到, 一开始情感在无聊的坐标点附近波动, 表示此时群体处于正常状态。随后情感轨迹突然移动到恐惧的坐标点附近, 说明群体情感转变为异常, 且变化十分突然。因此, 从定性的角度分析, 该实验对群体情感的描述符合事实情况。

[0139] 以上公开的仅为本发明的几个具体实施例, 但是, 本发明实施例并非局限于此, 任何本领域的技术人员能思之的变化都应落入本发明的保护范围。

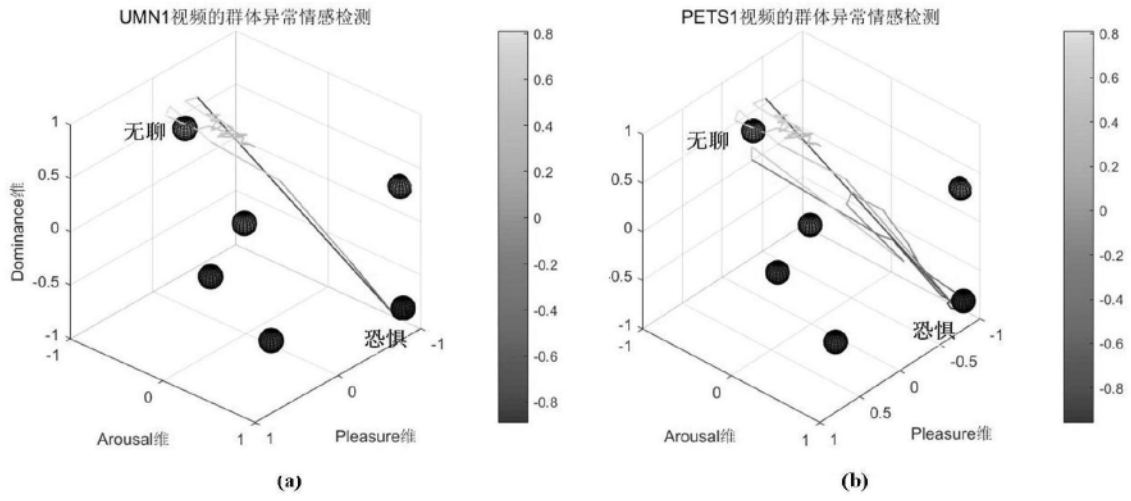


图1

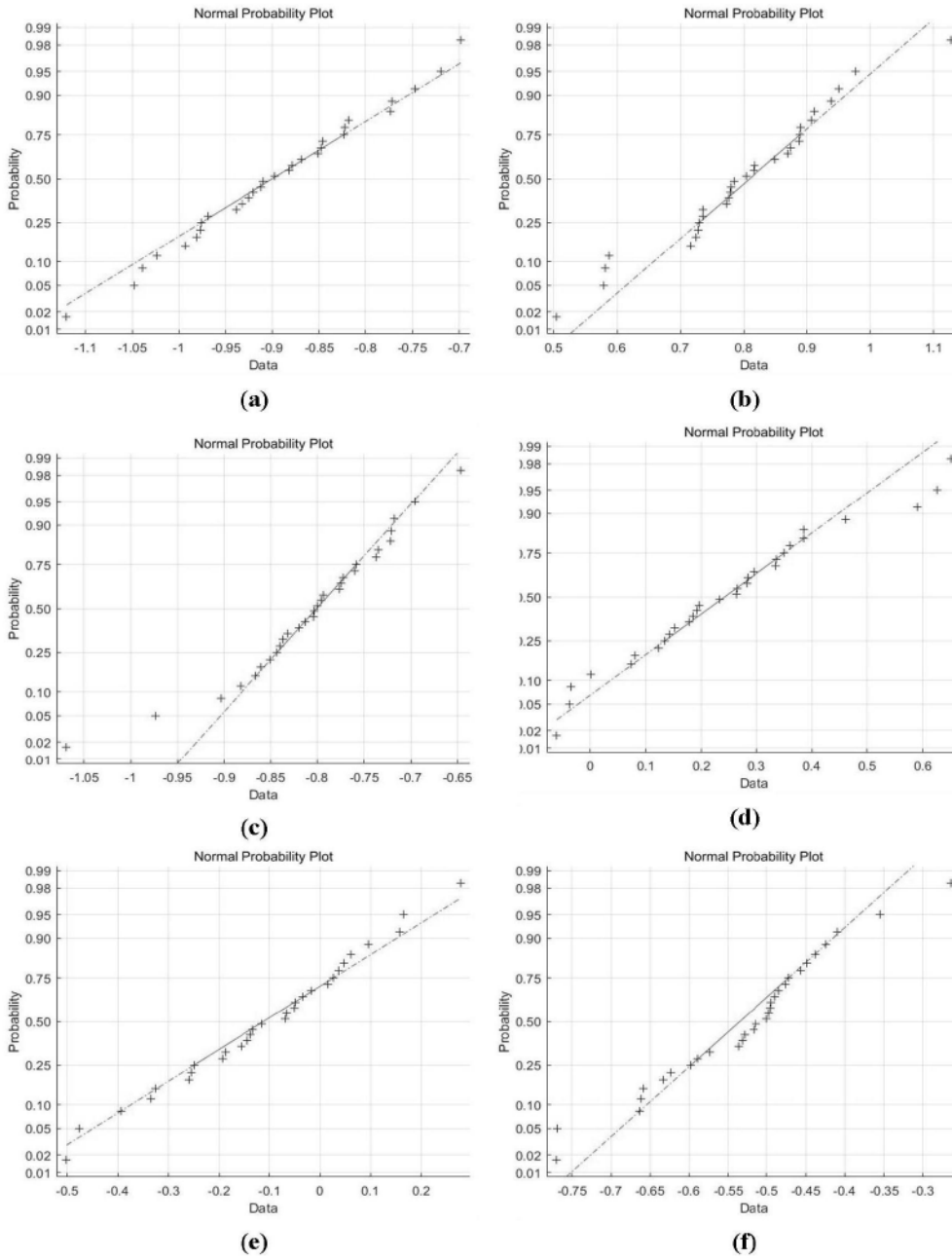


图2



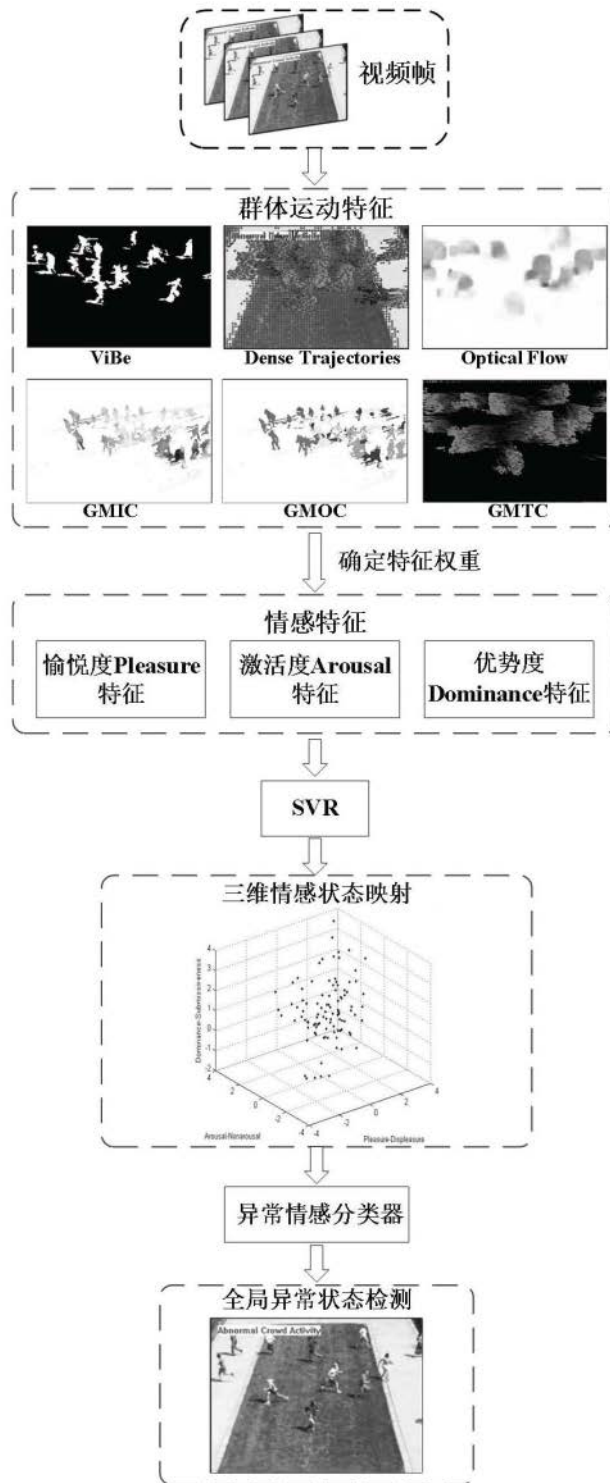


图3

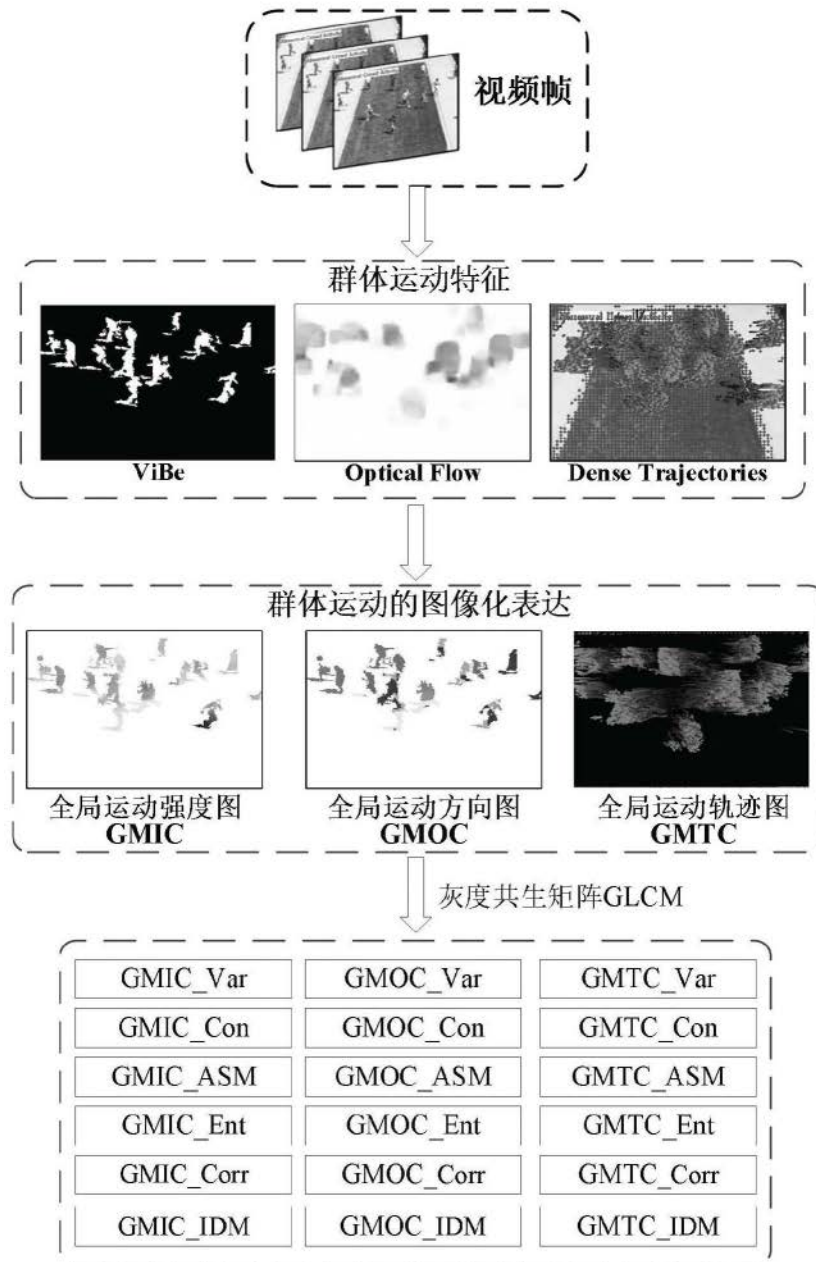


图4

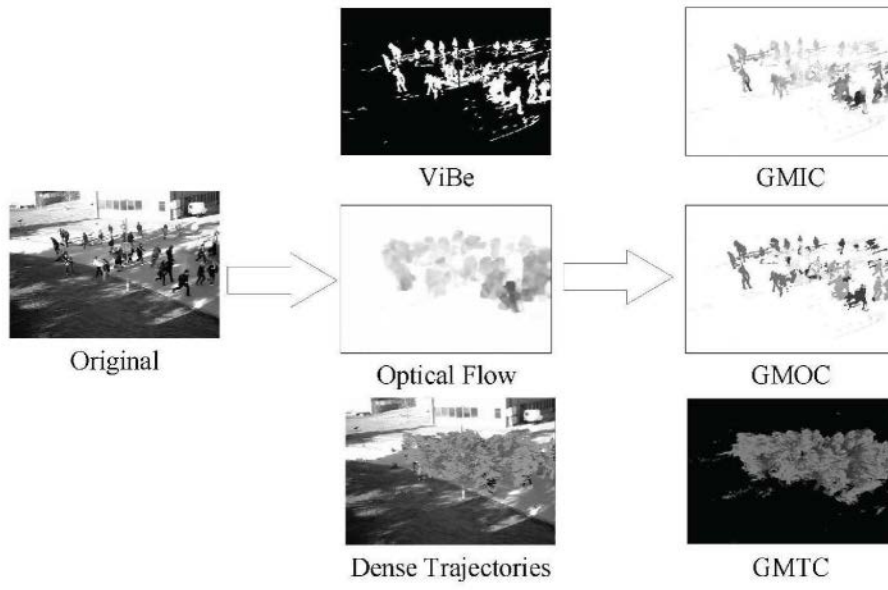


图5

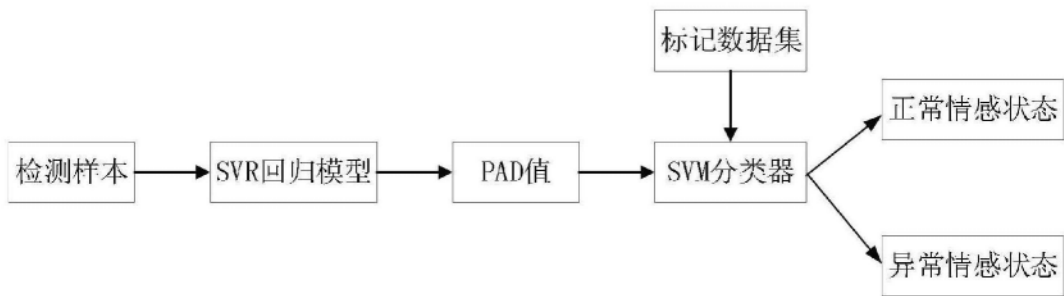


图6

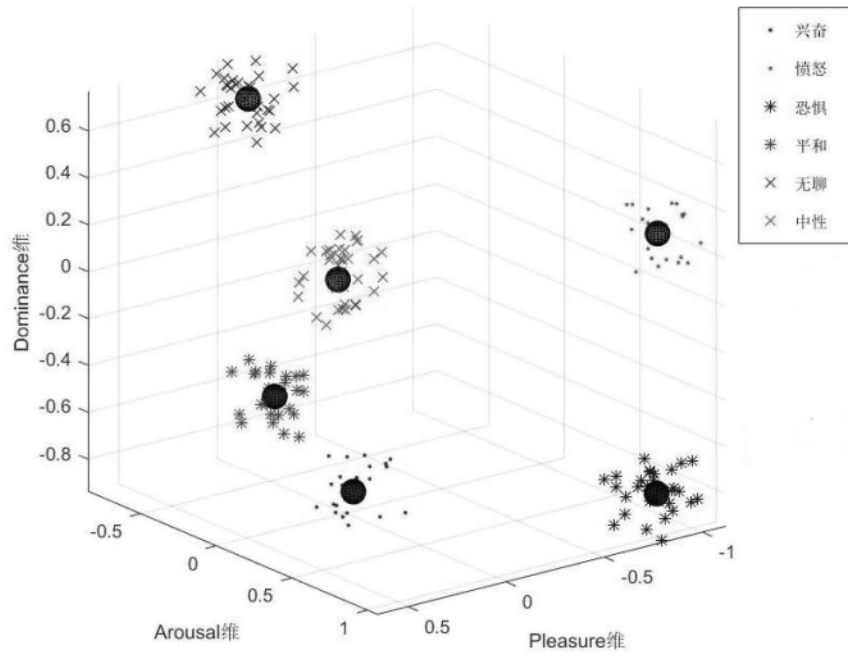


图7