



(12)发明专利申请

(10)申请公布号 CN 107239738 A

(43)申请公布日 2017. 10. 10

(21)申请号 201710310226.5

A61B 3/113(2006.01)

(22)申请日 2017.05.05

(71)申请人 南京邮电大学

地址 210023 江苏省南京市栖霞区栖霞街
道广月路9-1号

(72)发明人 叶宁 赵佳文 王娟 黄海平
王汝传 徐叶强 张力行 程康

(74)专利代理机构 南京知识律师事务所 32207
代理人 高玲玲

(51)Int.Cl.

G06K 9/00(2006.01)

G06F 17/18(2006.01)

A61B 5/16(2006.01)

A61B 5/024(2006.01)

A61B 5/00(2006.01)

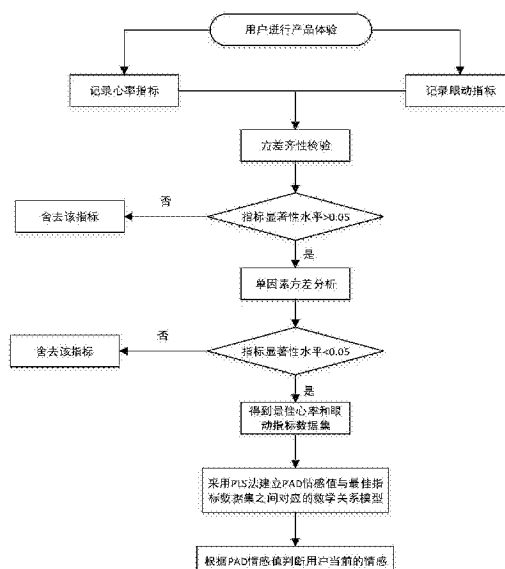
权利要求书2页 说明书7页 附图1页

(54)发明名称

一种融合眼动技术与心率检测技术的情感分析方法

(57)摘要

本发明公开了一种融合眼动技术与心率检测技术的情感分析方法,包括:步骤一:初步筛选出心率指标和眼动指标;步骤二:对用户产品体验过程中的相关心率和眼动指标数据进行记录,填写一份PAD情感调查问卷;步骤三:对所有初步筛选的心率和眼动指标数据进行方差齐性检测,进行筛选;步骤四:采用单因素方差分析法分别对心率和眼动指标数据进行分析,挑选得到与情感变化相关的最佳数据指标;步骤五:采用偏最小二乘回归分析法将得到的最佳心率和眼动指标与PAD情感值建立对应的数学关系模型;步骤六:根据PAD的情感值判断用户的情感状态。本发明融合两种不同的情感测量指标,这两种指标在情感识别过程中可以优势互补,能够有效识别不同的情感状态。



1. 一种融合眼动技术和心率检测技术的情感分析方法,其特征在于包含以下步骤:

S1: 用户接触产品A,亲身体验产品A的功能;

S2: 在用户对产品A进行体验的过程中,记录下相关心率和眼动指标数据,分别组成指标数据集B和C,并且在体验完后填写一份PAD情感调查问卷;

S3: 运用统计分析软件分别对B和C进行方差齐性检测,筛选出显著性水平大于特定值的指标组成集合B₁和C₁做进一步的情感分析;

S4: 采用单因素方差分析法分别对数据集B₁和C₁进行进一步分析,挑选出显著性水平小于0.05的指标,得到与情感变化相关的最佳数据指标集合B₂和C₂,将集合B₂和C₂合并成新的集合D作为自变量集合;

S5: 采用偏最小二乘回归分析法将得到的自变量集合D与PAD情感值,即因变量集合H,建立对应的数学关系模型;

S6: 由步骤5得到的PAD的情感值判断用户在产品体验过程中的情感状态。

2. 根据权利要求1所述的融合眼动技术和心率检测技术的情感分析方法,其特征在于步骤1中要求用户对产品A并不熟悉。

3. 根据权利要求1所述的融合眼动技术和心率检测技术的情感分析方法,其特征在于步骤3中所述特定值为0.05。

4. 根据权利要求1所述的融合眼动技术和心率检测技术的情感分析方法,其特征在于步骤4中所述单因素方差分析的具体步骤如下:

(1) 首先提出零假设,即观测变量各总体的均值在控制变量的不同水平下有无显著差异;

(2) 其次选择检验统计量: 方差分析采用的检验统计量是F统计量,即F值检验;

(3) 再次计算检验统计量的观测值和概率P值: 该步骤的目的就是计算检验统计量的观测值和相应的概率P值。

(4) 最后给定显著性水平,并作出决策: 如果相伴概率P小于显著性水平 α ,拒绝原假设,认为控制变量不同水平下各总体均值有显著差异,反之,则没有差异。

5. 根据权利要求1所述的融合眼动技术和心率检测技术的情感分析方法,其特征在于步骤5中所述偏最小二乘回归分析法的具体步骤如下:

(1) 首先对由自变量集合、因变量集合和样本集合构成的数据表 $X = (x_1, x_2, \dots, x_p)_{n \times p}$ 和 $Y = (y_1, y_2, \dots, y_q)_{n \times q}$ 进行标准化处理,分别得到两个矩阵为 $E_0 = (E_{01}, E_{02}, \dots, E_{0p})_{n \times p}$ 和 $F_0 = (F_{01}, F_{02}, \dots, F_{0q})_{n \times q}$ 并记 $\hat{t}_1$ 为 E_0 的第一个主成分;

(2) 其次求矩阵 $E_0^T F_0 F_0^T E_0$ 的最大特征值所对应的特征向量 w_1 , 计算成分得分向量 $\hat{t}_1 = E_0 w_1$;

(3) 再次进行交叉有效性检验,如果 $Q_h^2 < 0.0975$,则只取第一个主成分 $\hat{t}_1$ 就可以满足模型的精度要求,相反则需要进行下一步;

(4) 然后对 E_0 的残差矩阵 E_1 进行求解 $E_1 = E_0 - \hat{t}_1 \alpha_1^T$, 其中 $\alpha_1 = E_0^T \hat{t}_1 / \|\hat{t}_1\|^2$ 利用 E_1 和 F_0 重复第二步,直至达到条件 $Q_h^2 < 0.0975$,停止计算;

(5) 接着根据所提取的h个主成分 $(\hat{t}_1, \hat{t}_2, \dots, \hat{t}_h)$, 得到在 F_0 上的普通最小二乘回归模型:

$$F_0 = \hat{t}_1 \beta_1^T + \dots + \hat{t}_h \beta_h^T + F_h, \text{ 其中 } \beta_i = F_{i-1}^T \hat{t}_i / \|\hat{t}_i\|^2, \quad F_i = F_{i-1} - \hat{t}_i \beta_i^T \quad (i=1, 2, \dots, h);$$

(6) 最后把 $t_k^* = w_{k1}^* x_1 + \dots + w_{kp}^* x_p \quad (k=1, 2, \dots, h)$ 带入到 $Y = t_1 \beta_1 + \dots + t_h \beta_h$ 当中即得到q个因变量对应的偏最小二乘回归的方程式, 具体形式如下所示: $y_j = \alpha_{j1} x_1 + \dots + \alpha_{jp} x_p \quad (j=1, 2, \dots, q)$ 这里 $w_k^* = (w_{k1}^*, \dots, w_{kp}^*)$ 满足 $\hat{t}_k = E_0 w_k^*$, $w_k^* = \prod_{j=1}^{k-1} (1 - w_j \alpha_j^T) w_k$ 两项关系。

6. 根据权利要求5所述的融合眼动技术和心率检测技术的情感分析方法, 其特征在于所述交叉有效性检验的过程为: 在每次提取出新的成分 t_h 后, 从全部的样本集合中舍去第i个样本点 $(i=1, 2, \dots, n)$, 并使用之前的h主成分 $(t_1, t_2, \dots, t_h)$ 对已经去掉该样本点i之后的样本集合完成回归建模操作, 并最终形成回归方程, 然后把去掉的样本数据代入到该回归方程当中, 求得 $y_j \quad (j=1, 2, \dots, q)$ 在对应样本数据上的预测值 $\hat{y}_{(i)j}(h)$, 对 $i=1, 2, \dots, n$ 依次进行如上所示的计算步骤, 求得提取了h个主成分条件下的第j个因变量 $y_j \quad (j=1, 2, \dots, q)$ 对应的预测误差平方和为:

$$PRESS_j(h) = \sum_{i=1}^n (y_{ij} - \hat{y}_{(i)j}(h))^2 \quad (j=1, 2, \dots, q)$$

因变量的集合 $Y = (y_1, y_2, \dots, y_q)$ 所对应的预测误差的平方和为:

$$PRESS(h) = \sum_{j=1}^p PRESS_j(h) \quad (j=1, 2, \dots, q)$$

再采用所有的样本数据拟合含有h个主成分的对回归方程, 记 $\hat{y}_{ij}(h)$ 为样本点i的预测值, 则可以将因变量 y_j 的误差平方和定义为:

$$SS_j(h) = \sum_{i=1}^n (y_{ij} - \hat{y}_{ij}(h))^2 \quad (j=1, 2, \dots, q)$$

因变量的集合 $Y = (y_1, y_2, \dots, y_q)$ 的误差平方和为:

$$SS(h) = \sum_{j=1}^p SS_j(h) \quad (j=1, 2, \dots, q)$$

交叉有效性关系式为: $Q_h^2 = 1 - PRESS(h) / SS(h-1)$

在回归建模的每一步计算结束之前对模型进行交叉有效性检验, 判断条件为 $Q_h^2 < 1 - 0.95^2 = 0.0975$, 若在第h步满足条件, 则表明回归模型的精度已经达到要求, 不用再继续提取主成分; 相反若不满足条件, 则需要继续提取第 $(h+1)$ 个主成分。

7. 根据权利要求1所述的融合眼动技术和心率检测技术的情感分析方法, 其特征在于步骤5中所述PAD的情感值根据P、A、D三个维度值的正负关系可将情感划分为8种不同的类别: 1(+, +, +)表高兴, 2(-, -, -)表无聊, 3(+, +, -)表依赖, 4(-, -, +)表蔑视, 5(+, -, +)表放松, 6(-, +, -)表焦虑, 7(+, -, -)表温顺, 8(-, +, +)表敌意。

一种融合眼动技术与心率检测技术的情感分析方法

技术领域

[0001] 本发明属于情感认知计算领域,涉及一种融合眼动技术与心率检测技术的情感分析方法,采用生理指标(心率)检测技术和眼动技术相互融合的方法,建立心率、眼动指标与情感值之间的关系模型以实现情感的有效识别。

背景技术

[0002] 随着科学技术和计算机技术的快速发展,使得人类对计算机的依赖程度不断增强,人们对计算机的智能化能力要求越来越高。这种能力要求计算机能够像人一样思考,并且能够智能化地理解、表达人的情感,从而使计算机用户在一个和谐的人机交互环境中进行学习、工作和生活。

[0003] 随着用户体验经济的到来,人们对于产品的关注重点不再仅仅是产品的功能性、稳定性与安全性,而更多的是用户在使用产品过程中的满足感,也就是所谓的良好用户体验。用户体验在很大程度上受主观影响,在人机交互过程中不同的产品可以刺激出不同的情感体验,此时用户的情感状态最能代表其在使用过程中的真实感受。

[0004] 情感分析包括主观分析法和客观分析法,客观分析主要是通过对生理指标、眼动行为、语音特征等客观数据的分析来识别用户的情感特征。主观分析主要以自我报告或填写调查问卷的方式进行。目前,对于研究情感与能够反映其特征的客观数据的对应关系,我们大都选用单一的生理指标、语音信息、眼动数据来进行分析,因此情感的识别率可能比较低。人的生理信号主要是受人的自主神经系统和内分泌系统支配,而不受人的主观控制。如生理信号中的心率信号,当人的情感发生变化时,心率信号会产生不同程度的变化,尤其是在负面情感的情况下,心率信号的变化更为明显。眼睛是信息产品进行交互的第一窗口,也是用户感受产品特性的一个重要组成部分,且眼动数据是人机交互过程中伴随着情感变化而产生的客观准确的数据与信号。这两种信号都能以最直接最客观的形式表现出用户当下的情感状态。现有技术中目前还没有披露一种能够融合眼动技术和心率检测技术的情感分析,实现比较全面的进行情感识别的方法。

发明内容

[0005] 本发明的目的是提供一种融合眼动技术和心率检测技术的情感分析方法,该方法能够有效识别用户在产品体验过程中的情感状态。与很多已知的单一的情感识别相关方法不同,本发明是对多种情感测量指标进行统计分析,并结合PAD情感值进行融合建立新的情感识别模型的方法,通过使用本发明提出的新方法可以有效识别情感状态。

[0006] 为实现上述目的,本发明采用的技术方案为一种融合眼动技术和心率检测技术的情感分析方法,具体包含以下步骤:

[0007] S1:用户接触产品A,亲身体验产品A的功能;

[0008] S2:在用户对产品A进行体验的过程中,记录下相关心率和眼动指标数据,分别组成指标数据集B和C,并且在体验完后填写一份PAD情感调查问卷;

[0009] S3:运用统计分析软件分别对B和C进行方差齐性检测,筛选出显著性水平大于特定值的指标组成集合B₁和C₁做进一步的情感分析;

[0010] S4:采用单因素方差分析法分别对数据集B₁和C₁进行进一步分析,挑选出显著性水平小于0.05的指标,得到与情感变化相关的最佳数据指标集合B₂和C₂,将集合B₂和C₂合并成新的集合D作为自变量集合;

[0011] S5:采用偏最小二乘回归分析法将得到的自变量集合D与PAD情感值,即因变量集合H,建立对应的数学关系模型:

[0012] S6:由步骤5得到的PAD的情感值判断用户在产品体验过程中的情感状态。

[0013] 进一步,步骤1中要求用户对产品A并不熟悉。

[0014] 作为优选,步骤3中所述特定值为0.05。

[0015] 进一步,步骤4中所述单因素方差分析的具体步骤如下:

[0016] (1) 首先提出零假设,即观测变量各总体的均值在控制变量的不同水平下有无显著差异;

[0017] (2) 其次选择检验统计量:方差分析采用的检验统计量是F统计量,即F值检验;

[0018] (3) 再次计算检验统计量的观测值和概率P值:该步骤的目的就是计算检验统计量的观测值和相应的概率P值。

[0019] (4) 最后给定显著性水平,并作出决策:如果相伴概率P小于显著性水平 α ,拒绝原假设,认为控制变量不同水平下各总体均值有显著差异,反之,则没有差异。

[0020] 进一步,步骤5中所述偏最小二乘回归分析法的具体步骤如下:

[0021] (1) 首先对由自变量集合、因变量集合和样本集合构成的数据表 $X = (x_1, x_2, \dots, x_p)_{n \times p}$ 和 $Y = (y_1, y_2, \dots, y_q)_{n \times q}$ 进行标准化处理,分别得到两个矩阵为 $E_0 = (E_{01}, E_{02}, \dots, E_{0p})_{n \times p}$ 和 $F_0 = (F_{01}, F_{02}, \dots, F_{0q})_{n \times q}$ 并记 $\hat{t}_1$ 为 E_0 的第一个主成分;

[0022] (2) 其次求矩阵 $E_0^T F_0 F_0^T E_0$ 的最大特征值所对应的特征向量 w_1 ,计算成分得分向量 $\hat{t}_1 = E_0 w_1$;

[0023] (3) 再次进行交叉有效性检验,如果 $Q_h^2 < 0.0975$,则只取第一个主成分 $\hat{t}_1$ 就可以满足模型的精度要求,相反则需要进行下一步;

[0024] (4) 然后对 E_0 的残差矩阵 E_1 进行求解 $E_1 = E_0 - \hat{t}_1 \alpha_1^T$,其中 $\alpha_1 = E_0^T \hat{t}_1 / \|\hat{t}_1\|^2$ 利用 E_1 和 F_0 重复第二步,直至达到条件 $Q_h^2 < 0.0975$,停止计算;

[0025] (5) 接着根据所提取的h个主成分 $(\hat{t}_1, \hat{t}_2, \dots, \hat{t}_h)$,得到在 F_0 上的普通最小二乘回归模型: $F_0 = \hat{t}_1 \beta_1^T + \dots + \hat{t}_h \beta_h^T + F_h$,其中 $\beta_i = F_{i-1}^T \hat{t}_i / \|\hat{t}_i\|^2$, $F_i = F_{i-1} - \hat{t}_i \beta_i^T$ ($i=1, 2, \dots, h$);

[0026] (6) 最后把 $t_k = w_{k1}^* x_1 + \dots + w_{kp}^* x_p$ ($k=1, 2, \dots, h$)带入到 $Y = t_1 \beta_1 + \dots + t_h \beta_h$ 当中即得到q个因变量对应的偏最小二乘回归的方程式,具体形式如下所示: $y_j = \alpha_{j1} x_1 + \dots + \alpha_{jp} x_p$ ($j=1, 2, \dots, q$)

[0027] 这里 $w_k^* = (w_{k1}^*, \dots, w_{kp}^*)$ 满足 $\hat{t}_k = E_0 w_k^*$, $w_k^* = \prod_{j=1}^{k-1} (1 - w_j \alpha_j^T) w_k$ 两项关系。

[0028] 上述交叉有效性检验检验的过程为:在每次提取出新的成分 t_h 后,从全部的样本

集合中舍去第*i*个样本点 ($i=1,2,\dots,n$), 并使用之前的*h*主成分 ($t_1, t_2, \dots, t_h$) 对已经去掉该样本点*i*之后的样本集合完成回归建模操作, 并最终形成回归方程, 然后把去掉的样本数据代入到该回归方程当中, 求得 y_j ($j=1,2,\dots,q$) 在对应样本数据上的预测值 $\hat{y}_{(i)j}(h)$, 对*i*=1,2,...,n依次进行如上所示的计算步骤, 求得提取了*h*个主成分条件下的第*j*个因变量 y_j ($j=1,2,\dots,q$) 对应的预测误差平方和为:

$$[0029] \quad PRESS_j(h) = \sum_{i=1}^n (y_{ij} - \hat{y}_{(i)j}(h))^2 \quad (j=1,2,\dots,q)$$

[0030] 因变量的集合 $Y = (y_1, y_2, \dots, y_q)$ 所对应的预测误差的平方和为:

$$[0031] \quad PRESS(h) = \sum_{j=1}^q PRESS_j(h) \quad (j=1,2,\dots,q)$$

[0032] 再采用所有的样本数据拟合含有*h*个主成分的对应回归方程, 记 $\hat{y}_{ij}(h)y_j$ 为样本点*i*的预测值, 则可以将因变量 y_j 的误差平方和定义为:

$$[0033] \quad SS_j(h) = \sum_{i=1}^n (y_{ij} - \hat{y}_{ij}(h))^2 \quad (j=1,2,\dots,q)$$

[0034] 因变量的集合 $Y = (y_1, y_2, \dots, y_q)$ 的误差平方和为:

$$[0035] \quad SS(h) = \sum_{j=1}^q SS_j(h) \quad (j=1,2,\dots,q)$$

[0036] 交叉有效性关系式为: $Q_h^2 = 1 - PRESS(h) / SS(h-1)$

[0037] 在回归建模的每一步计算结束之前对模型进行交叉有效性检验, 判断条件为 $Q_h^2 < 1 - 0.95^2 = 0.0975$, 若在第*h*步满足条件, 则表明回归模型的精度已经达到要求, 不用再继续提取主成分; 相反若不满足条件, 则需要继续提取第(*h*+1)个主成分。

[0038] 进一步, 上述步骤5中所述PAD的情感值根据P、A、D三个维度值的正负关系可将情感划分为8种不同的类别: 1(+、+、+)表高兴, 2(-、-、-)表无聊, 3(+、+、-)表依赖, 4(-、-、+)表蔑视, 5(+、-、+)表放松, 6(-、+、-)表焦虑, 7(+、-、-)表温顺, 8(-、+、+)表敌意。

[0039] 与现有技术相比, 本发明的有益效果:

[0040] (1) 本发明提出的方法突破了单一情感指标与情感之间关系的研究局限, 融合两种不同的情感测量指标, 这两种指标在情感识别过程中可以优势互补, 能够有效识别不同的情感状态;

[0041] (2) 本发明提出的方法是基于眼动和心率指标的分析, 这两种指标都是伴随情感变化而产生的客观数据, 是用户的情感状态的真实反映;

[0042] (3) 本发明采用两种统计分析方法来寻找情感与眼动和心率指标之间的对应关系, 简单易懂, 易于实现;

附图说明

[0043] 图1是本发明的流程示意图。

具体实施方式

[0044] 现结合附图对本发明作进一步详细的说明。本发明提出的一种融合眼动技术和心率检测技术的情感分析方法, 如图1所示, 主要包括: 用户进行产品体验, 结束后填写一份

PAD情感问卷;对用户在产品体验过程中产生的眼动指标和心率指标进行方差齐性检验和单因素方差分析,筛选出与情感变化有关的最佳的眼动和心率指标;通过偏最小二乘法将这两种指标与情感值建立对应的数学关系模型。以用户体验产品A为例,包含以下步骤:

[0045] 步骤1:用户接触产品A(为确保效果,要求用户对产品A并不熟悉),并进行亲身体验产品A的功能;

[0046] 步骤2:在用户对产品A进行使用体验的过程中,记录下相关心率和眼动指标数据,分别组成指标数据集B和C,并且在体验完后填写一份PAD情感调查问卷;

[0047] 步骤3:运用SPSS软件分别对所有初步筛选的心率和眼动指标数据集B和进行方差齐性检测(方法1),筛选出显著性水平大于0.05的指标组成集合B₁和C₁中,做进一步的情感分析;

[0048] 步骤4:采用单因素方差分析法(方法2)分别对满足方差齐性的心率和眼动指标数据集B₁和C₁进行进一步分析,挑选出显著性水平小于0.05的指标,得到与情感变化相关的最佳数据指标集合B₂和C₂,将集合B₂和C₂合并成新的集合D作为自变量集合;

[0049] 步骤5:采用偏最小二乘回归分析法(PLS)(方法3)将得到的最佳心率和眼动指标自变量集合D与PAD情感值(因变量集合H)建立对应的数学关系模型:

[0050] (1) 首先对由自变量集合(假设集合D中元素有p个)、因变量集合(集合H包括P、A、D三个情感值元素)和样本集合(假设元素有n个)构成的数据表 $X = (x_1, x_2, \dots, x_p)_{n \times p}$ 和 $Y = (y_1, y_2, y_3)_{n \times 3}$ 进行标准化处理,分别得到两个矩阵为 $E_0 = (E_{01}, E_{02}, \dots, E_{0p})_{n \times p}$ 和 $F_0 = (F_{01}, F_{02}, F_{03})_{n \times 3}$ 并记 $\hat{t}_1$ 为 E_0 的第一个主成分。

[0051] (2) 其次求矩阵 $E_0^T F_0 F_0^T E_0$ 的最大特征值所对应的特征向量 w_1 ,计算成分得分向量 $\hat{t}_1 = E_0 w_1$ 。

[0052] (3) 再次进行交叉有效性检验,如果 $Q_h^2 < 0.0975$,则只取第一个主成分 $\hat{t}_1$ 就可以满足模型的精度要求,相反则需要进行下一步。

[0053] (4) 然后对 E_0 的残差矩阵 E_1 进行求解 $E_1 = E_0 - \hat{t}_1 \alpha_1^T$,其中 $\alpha_1 = E_0^T \hat{t}_1 / \|\hat{t}_1\|^2$ 利用 E_1 和 F_0 重复第二步,直至达到条件 $Q_h^2 < 0.0975$,停止计算。

[0054] (5) 接着根据所提取的h个主成分 $(\hat{t}_1, \hat{t}_2, \dots, \hat{t}_h)$,得到在 F_0 上的普通最小二乘回归模型: $F_0 = \hat{t}_1 \beta_1^T + \dots + \hat{t}_h \beta_h^T + F_h$,其中 $\beta_i = F_{i-1}^T \hat{t}_i / \|\hat{t}_i\|^2$, $F_i = F_{i-1} - \hat{t}_i \beta_i^T$ ($i=1, 2, \dots, h$)。

[0055] (6) 最后把 $t_k = w_{k1}^* x_1 + \dots + w_{kp}^* x_p$ ($k=1, 2, \dots, h$)带入到 $Y = t_1 \beta_1 + \dots + t_h \beta_h$ 当中即得到3个因变量对应的偏最小二乘回归的方程式,具体形式如下所示: $y_j = \alpha_{j1} x_1 + \dots + \alpha_{jp} x_p$ ($j=1, 2, 3$)

[0056] 这里 $w_k^* = (w_{k1}^* \dots w_{kp}^*)$ 满足 $\hat{t}_k = E_0 w_k^*$, $w_k^* = \prod_{j=1}^{k-1} (1 - w_j \alpha_j^T) w_k$ 两项关系。

[0057] 步骤6:由步骤五得到的PAD的情感值(y_j)判断用户在产品体验过程中的情感状态。根据P、A、D三个维度值的正负关系可将情感划分为8种不同的类别:1(+、+、+)表高兴,2(-、-、-)表无聊,3(+、+、-)表依赖,4(-、-、+)表蔑视,5(+、-、+)表放松,6(-、+、-)表焦虑,7(+、-、-)表温顺,8(-、+、+)表敌意。

[0058] 现对本发明中所用的方法进行说明：

[0059] 方法1：方差齐性检验，是对控制变量不同水平下各观测变量总体方差是否相等进行检验。单因素方差分析中，方差齐性检验采用了方差同质性 (Homogeneity of Variance) 检验方法，其原假设是：各水平下观测变量总体的方差无显著差异。因此，在单因素方差分析前有必要对方差是否齐性进行检验，若其显著性水平值大于0.05，满足方差整齐 (即方差无显著差异) 的假设，则观测变量可用于进行下一步的单因素方差分析。

[0060] 方法2：单因素方差分析是一种用于研究观测变量在单一控制变量的影响下是否有显著性差异的方法，其基本思想是：先确定观测变量以及对应的控制变量；再分析观测变量的方差，在方差分析中观测变量值的变动将受到控制变量和随机变量两方面的影响；最后比较观测变量总离差平方和各部分所占的比例，推断控制变量是否给观测变量带来了显著影响。具体步骤如下：

[0061] (5) 首先提出零假设，即观测变量各总体的均值在控制变量的不同水平下有无显著差异。

[0062] (6) 其次选择检验统计量：方差分析采用的检验统计量是F统计量，即F值检验。

[0063] (7) 再次计算检验统计量的观测值和概率P值：该步骤的目的就是计算检验统计量的观测值和相应的概率P值。

[0064] (8) 最后给定显著性水平，并作出决策：如果相伴概率P小于显著性水平 α ($\alpha = 0.05$)，拒绝原假设，认为控制变量不同水平下各总体均值有显著差异，反之，则没有差异。

[0065] 方法3：偏最小二乘法回归分析 (PLS) 就是考虑 q 个因变量 $y_1, y_2, \dots, y_q$ 与 p 个自变量 $x_1, x_2, \dots, x_p$ 的建模问题。基本思想是将自变量集合和因变量集合以及样本量为 n 的样本集合共同组成数据表 $X = (x_1, x_2, \dots, x_p)_{n \times p}$ 和 $Y = (y_1, y_2, \dots, y_q)_{n \times q}$ ，运用偏最小二乘法分别在自变量集合和因变量集合中提出主成分分别为 t_1 和 u_1 (t_1 是 $x_1, x_2, \dots, x_p$ 的线性组合， u_1 是 $y_1, y_2, \dots, y_q$ 的线性组合)， t_1 和 u_1 要求各自尽可能多地提取所在变量集合中的变异信息，且两者相关程度达到最大。然后建立因变量 $y_1, y_2, \dots, y_q$ 与 t_1 的回归，如果回归方程已达到满意的精度，则算法终止。否则继续第二对成分的提取，直到能达到满意的精度为止。若最终对自变量集提取 h 个成分 $t_1, t_2, \dots, t_h$ ，偏最小二乘回归将通过建立 $y_1, y_2, \dots, y_q$ 与 $t_1, t_2, \dots, t_h$ 的回归式，然后再表示为 $y_1, y_2, \dots, y_q$ 与原自变量的回归方程式，即偏最小二乘回归方程式。具体步骤如下：

[0066] (7) 首先对由自变量集合、因变量集合和样本集合构成的数据表 $X = (x_1, x_2, \dots, x_p)_{n \times p}$ 和 $Y = (y_1, y_2, \dots, y_q)_{n \times q}$ 进行标准化处理，分别得到两个矩阵为 $E_0 = (E_{01}, E_{02}, \dots, E_{0p})_{n \times p}$ 和 $F_0 = (F_{01}, F_{02}, \dots, F_{0q})_{n \times q}$ 并记 $\hat{t}_1$ 为 E_0 的第一个主成分。

[0067] (8) 其次求矩阵 $E_0^T F_0 F_0^T E_0$ 的最大特征值所对应的特征向量 w_1 ，计算成分得分向量 $\hat{t}_1 = E_0 w_1$ 。

[0068] (9) 再次进行交叉有效性检验，如果 $Q_h^2 < 0.0975$ ，则只取第一个主成分 $\hat{t}_1$ 就可以满足模型的精度要求，相反则需要进行下一步。

[0069] (10) 然后对 E_0 的残差矩阵 E_1 进行求解 $E_1 = E_0 - \hat{t}_1 \alpha_1^T$ ，其中 $\alpha_1 = E_0^T \hat{t}_1 / \|\hat{t}_1\|^2$ 利用 E_1 和 F_0 重复第二步，直至达到条件 $Q_h^2 < 0.0975$ ，停止计算。

[0070] (11) 接着根据所提取的h个主成分 $(\hat{t}_1, \hat{t}_2, \dots, \hat{t}_h)$, 得到在 F_0 上的普通最小二乘回归模型: $F_0 = \hat{t}_1 \beta_1^T + \dots + \hat{t}_h \beta_h^T + F_h$, 其中 $\beta_i = F_{i-1}^T \hat{t}_i / \|\hat{t}_i\|^2$, $F_i = F_{i-1} - \hat{t}_i \beta_i^T$ ($i=1, 2, \dots, h$)。

[0071] (12) 最后把 $\hat{t}_k = w_{k1}^* x_1 + \dots + w_{kp}^* x_p$ ($k=1, 2, \dots, h$) 带入到 $Y = t_1 \beta_1 + \dots + t_h \beta_h$ 当中即得到q个因变量对应的偏最小二乘回归的方程式, 具体形式如下所示: $y_j = \alpha_{j1} x_1 + \dots + \alpha_{jp} x_p$ ($j=1, 2, \dots, q$)

[0072] 这里 $w_k^* = (w_{k1}^* \dots w_{kp}^*)$ 满足 $\hat{t}_k = E_0 w_k^*$, $w_k^* = \prod_{j=1}^{k-1} (1 - w_j \alpha_j^T) w_k$ 两项关系。

[0073] 方法4: 交叉有效性检验, 一般情况下, 偏最小二乘法不需要选用存在的h个成分 $t_1, t_2, \dots, t_h$ 来建立回归式, 而像主成分分析一样, 只选用前1个成分 ($1 \leq h$), 就可以得到预测能力较好的回归模型。对于建模所需要提取的成分个数1, 可以通过交叉有效性检验来确定。

[0074] 检验过程在每次提取出新的成分 t_h 后, 从全部的样本集合中舍去第i个样本点 ($i=1, 2, \dots, n$), 并使用之前的h主成分 ($t_1, t_2, \dots, t_h$) 对已经去掉该样本点i之后的样本集合完成回归建模操作, 并最终形成回归方程, 然后把去掉的样本数据代入到该回归方程当中, 求得 y_j ($j=1, 2, \dots, q$) 在对应样本数据上的预测值 $\hat{y}_{(i)j}(h)$, 对 $i=1, 2, \dots, n$ 依次进行如上所示的计算步骤, 求得提取了h个主成分条件下的第j个因变量 y_j ($j=1, 2, \dots, q$) 对应的预测误差平方和为:

$$[0075] \quad PRESS_j(h) = \sum_{i=1}^n (y_{ij} - \hat{y}_{(i)j}(h))^2 \quad (j=1, 2, \dots, q)$$

[0076] 因变量的集合 $Y = (y_1, y_2, \dots, y_q)$ 所对应的预测误差的平方和为:

$$[0077] \quad PRESS(h) = \sum_{j=1}^q PRESS_j(h) \quad (j=1, 2, \dots, q)$$

[0078] 另外, 再采用所有的样本数据拟合含有h个主成分的对应回归方程。记 $\hat{y}_{ij}(h) y_j$ 为样本点i的预测值, 则可以将因变量 y_j 的误差平方和定义为:

$$[0079] \quad SS_j(h) = \sum_{i=1}^n (y_{ij} - \hat{y}_{ij}(h))^2 \quad (j=1, 2, \dots, q)$$

[0080] 因变量的集合 $Y = (y_1, y_2, \dots, y_q)$ 的误差平方和为:

$$[0081] \quad SS(h) = \sum_{j=1}^q SS_j(h) \quad (j=1, 2, \dots, q)$$

[0082] 当变量 $PRESS(h)$ 到达其最小状态时, 此时的h值就是要求的主成分个数。一般有关系 $PRESS(h) > SS(h)$, $SS(h) < SS(h-1)$ 。因此在主成分提取的过程中, $PRESS(h) / SS(h-1)$ 的比值状态为越小越好, 因此在此过程中将其设定为0.05, 此处可以理解为当满足条件 $PRESS(h) / SS(h-1) \leq (1-0.05)^2 = 0.95^2$ 时, 可以通过增加主成分达到使模型精确度明显提高的目的。综上所述交叉有效性关系式为:

$$[0083] \quad Q_h^2 = 1 - PRESS(h) / SS(h-1)$$

[0084] 按照这种方法需要在回归建模的每一步计算结束之前对模型进行交叉有效性检验, 判断条件为 $Q_h^2 < 1 - 0.95^2 = 0.0975$, 若在第h步满足条件, 则表明回归模型的精度已经达到要求, 不用再继续提取主成分; 相反若不满足条件, 则需要继续提取第(h+1)个主成分。

[0085] 为便于本领域的普通技术人员理解和实施本发明的方法,现提出一个具体的实施例,如下:

[0086] 步骤1:通过某大学硕士生徐某对心率信号的各项指标与情感之间对应关系的研究发现,一些心率指标如:心率上升沿幅度的方差、心率上升沿幅度的最小值、心率上升沿所用时间的最小值、心率上升斜率大小的最大值、心率幅度相邻差值的均值、心率下降斜率大小的方差等在情感识别中发挥着较大的作用,以及另一所大学的孙某对眼动技术的研究发现眼动指标如:眨眼率、平均眨眼时间、最大眨眼时间、最小眨眼时间、注视率、扫视率、最大扫视时间、扫视平均幅度、扫视最大幅度、扫视平均速度、瞳孔最大值和瞳孔平均值等能较好的服务于情感识别。因此,初步筛选出以上心率和眼动指标用于情感分析;

[0087] 步骤2:对用户产品体验过程中的相关心率和眼动指标数据进行记录,并且在体验完后填写一份PAD情感调查问卷;

[0088] 步骤3:运用SPSS软件对所有初步筛选的心率和眼动指标数据进行方差齐性检测(方法1),筛选出显著性水平大于0.05的指标,做进一步的情感分析;

[0089] 步骤4:采用单因素方差分析法(方法2)分别对满足方差齐性的心率和眼动指标数据进行进一步分析,挑选出显著性水平小于0.05的指标,得到与情感变化相关的最佳数据指标;

[0090] 步骤5:采用偏最小二乘回归分析法(PLS)(方法3)将得到的最佳心率和眼动指标与PAD情感值建立对应的数学关系模型;

[0091] 步骤6:根据PAD的情感值判断用户在产品体验过程中的情感状态。

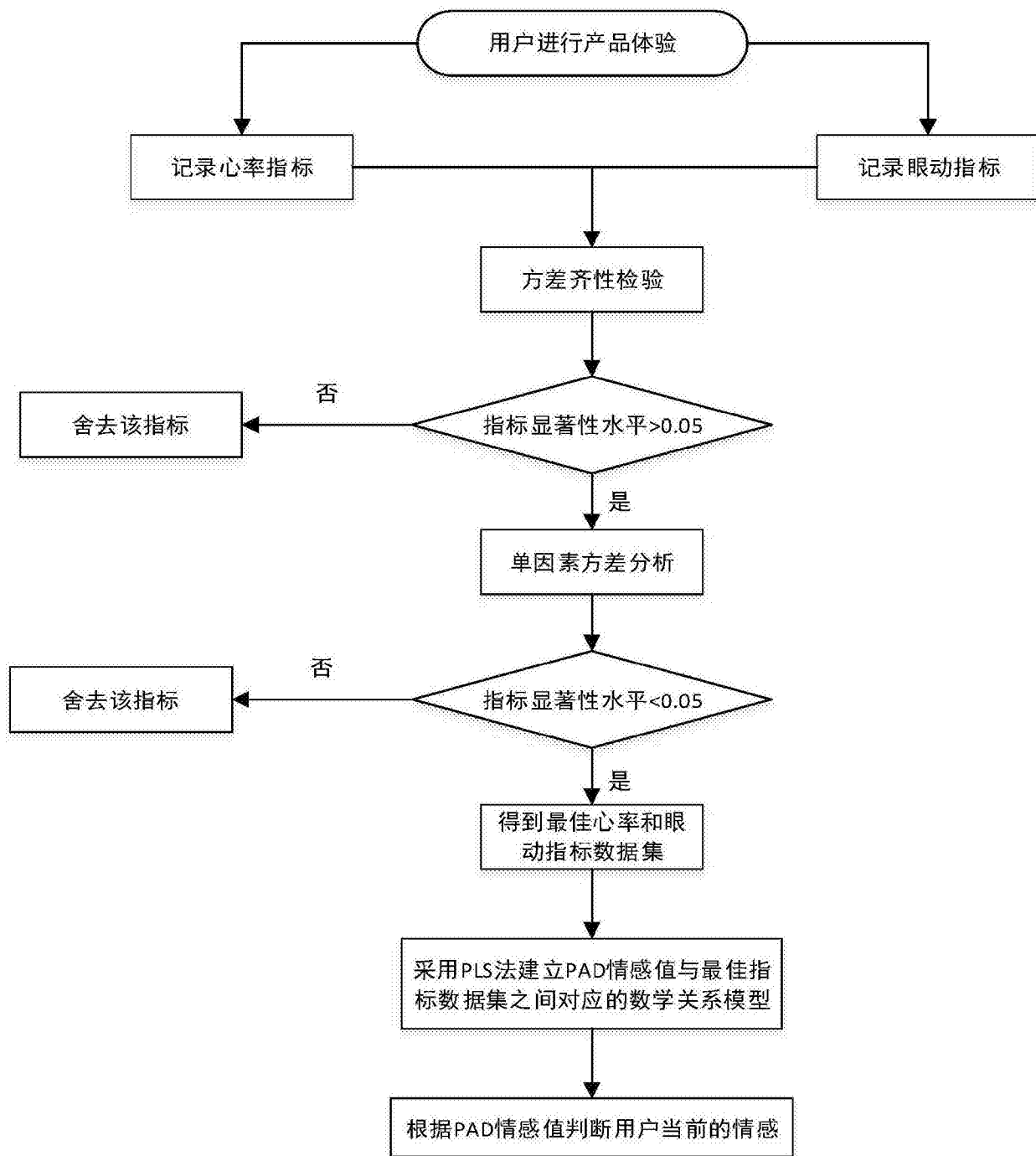


图1

专利名称(译)	一种融合眼动技术与心率检测技术的情感分析方法		
公开(公告)号	CN107239738A	公开(公告)日	2017-10-10
申请号	CN2017110310226.5	申请日	2017-05-05
[标]申请(专利权)人(译)	南京邮电大学		
申请(专利权)人(译)	南京邮电大学		
当前申请(专利权)人(译)	南京邮电大学		
[标]发明人	叶宁 赵佳文 王娟 黄海平 王汝传 徐叶强 张力行 程康		
发明人	叶宁 赵佳文 王娟 黄海平 王汝传 徐叶强 张力行 程康		
IPC分类号	G06K9/00 G06F17/18 A61B5/16 A61B5/024 A61B5/00 A61B3/113		
CPC分类号	A61B3/113 A61B5/024 A61B5/165 A61B5/72 G06F17/18 G06K9/00885 G06K2009/00939		
代理人(译)	高玲玲		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了一种融合眼动技术与心率检测技术的情感分析方法，包括：步骤一：初步筛选出心率指标和眼动指标；步骤二：对用户产品体验过程中的相关心率和眼动指标数据进行记录，填写一份PAD情感调查问卷；步骤三：对所有初步筛选的心率和眼动指标数据进行方差齐性检测，进行筛选；步骤四：采用单因素方差分析法分别对心率和眼动指标数据进行分析，挑选得到与情感变化相关的最佳数据指标；步骤五：采用偏最小二乘回归分析法将得到的最佳心率和眼动指标与PAD情感值建立对应的数学关系模型；步骤六：根据PAD的情感值判断用户的情感状态。本发明融合两种不同的情感测量指标，这两种指标在情感识别过程中可以优势互补，能够有效识别不同的情感状态。

