



(12) 发明专利申请

(10) 申请公布号 CN 115830682 A

(43) 申请公布日 2023. 03. 21

(21) 申请号 202211562229.5

(22) 申请日 2022.12.07

(71) 申请人 东南大学

地址 211102 江苏省南京市江宁区东南大学路2号

(72) 发明人 郑文明 魏梦婷 江星洵 宗源路成

(74) 专利代理机构 南京苏高专利商标事务所
(普通合伙) 32204

专利代理师 冯艳芬

(51) Int. Cl.

G06V 40/16 (2022.01)

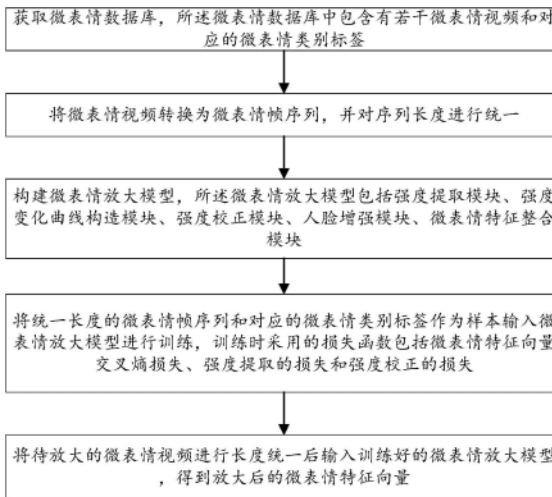
权利要求书3页 说明书8页 附图2页

(54) 发明名称

样本自适应的微表情放大方法及装置

(57) 摘要

本发明公开了一种基于样本自适应的微表情放大方法及装置,包括:(1)获取微表情数据库;(2)将微表情视频转换为微表情帧序列,并对序列长度进行统一;(3)构建微表情放大模型,所述微表情放大模型包括:强度提取模块,强度变化曲线构造模块,强度校正模块,人脸增强模块,微表情特征整合模块,(4)将统一长度的微表情帧序列和对应的微表情类别标签作为样本输入微表情放大模型进行训练,训练时采用的损失函数包括微表情特征向量交叉熵损失、强度提取的损失和强度校正的损失;(5)将待放大的微表情视频进行长度统一后输入训练好的微表情放大模型,得到放大后的微表情特征向量。本发明可以根据每个微表情视频进行自适应放大,效果更好。



1. 一种基于样本自适应的微表情放大方法,其特征在于该方法包括:

(1) 获取微表情数据库,所述微表情数据库中包含有若干微表情视频和对应的微表情类别标签;

(2) 将微表情视频转换为微表情帧序列,并对序列长度进行统一;

(3) 构建微表情放大模型,所述微表情放大模型包括:

强度提取模块,用于利用微表情帧序列之间的强度差异,通过对比学习的方式提取强度值;

强度变化曲线构造模块,用于基于高斯分布构造微表情帧序列的强度变化曲线;

强度校正模块,用于采用秩和检验的方式,将所述强度提取模块提取到的强度值按照所述强度变化曲线的趋势进行校正;

人脸增强模块,用于将校正后的强度值与从微表情帧序列中提取到的人脸特征整合,得到增强人脸特征向量;

微表情特征整合模块,用于将属于一个微表情视频的所有帧的增强人脸特征向量使用LSTM整合,作为放大后的微表情特征向量;

(4) 将统一长度的微表情帧序列和对应的微表情类别标签作为样本输入微表情放大模型进行训练,训练时采用的损失函数包括微表情特征向量交叉熵损失、强度提取的损失和强度校正的损失;

(5) 将待放大的微表情视频进行长度统一后输入训练好的微表情放大模型,得到放大后的微表情特征向量。

2. 根据权利要求1所述的样本自适应的微表情放大方法,其特征在于:步骤(2)中所述对序列长度进行统一具体包括:

从每个微表情帧序列中按照预设间隔抽取N帧,形成长度都为N的微表情帧序列,N为正整数。

3. 根据权利要求1所述的样本自适应的微表情放大方法,其特征在于:所述利用微表情帧序列之间的强度差异,通过对比学习的方式提取强度值,具体包括:

(3-1-1) 对于统一后的每个微表情帧序列,将每帧x作为源样本,对x使用数据增强生成两个新的图像 \tilde{x}_1 和 \tilde{x}_2 ,作为x的正样本,从序列除x外的剩余帧中选择M张图片,记为 $x_1^-, x_m^-, \dots, x_M^-$,作为x的负样本,M为小于N的正整数,N表示微表情帧序列的长度;

(3-1-2) 将每帧的正样本 \tilde{x}_1, \tilde{x}_2 和负样本 $x_1^-, x_m^-, \dots, x_M^-$ 依次输入Resnet网络+FC线性层的网络中,得到特征向量 \tilde{z}_1, \tilde{z}_2 和 $z_1^-, z_m^-, \dots, z_M^-$;

(3-1-3) 将每帧的特征向量 $\tilde{z}_1, \tilde{z}_2, z_1^-, z_m^-, \dots, z_M^-$ 输入一线性层,得到对应的强度值;

(3-1-4) 采用以下公式计算强度提取的损失,从而实现训练过程中强度信息提取的优化:

$$\mathcal{L}_{ID} = \sum_{k \in D} \frac{-1}{N} \sum_{i=1}^N \left[\log \frac{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i})}{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}) + \sum_{m=1}^M \text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_m^{-k,i})} + \log \frac{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i})}{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}) + \sum_{m=1}^M \text{sim}(\tilde{z}_2^{k,i}, \tilde{z}_m^{-k,i})} \right]$$

式中, i 表示帧索引, D 表示微表情数据库的微表情视频数, k 表示微表情视频的索引, $\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}, \tilde{z}_m^{-k,i}$ 分别表示微表情数据库中第 k 个微表情视频长度统一后的第 i 帧的特征向量 $\tilde{z}_1, \tilde{z}_2, \tilde{z}_m$ 。

4. 根据权利要求1所述的样本自适应的微表情放大方法, 其特征在于: 所述基于高斯分布构造微表情帧序列的强度变化曲线具体包括:

(3-2-1) 设定微表情帧序列强度的变化范围在 $[\epsilon, 1]$, 其中 $\epsilon > 0$;

(3-2-2) 以峰值帧为界限将微表情帧序列分为两段, 从起始帧到峰值帧记为 $0n-A$ 段, 从峰值帧到结束帧记为 $A-Off$ 段, 其中 $0n-A$ 段含有 N_1 帧, $A-Off$ 段含有 N_2 帧, $N_1 + N_2 + 1 = N$, N 表示微表情帧序列的长度;

(3-2-3) 设置 $0n-A$ 段上每帧的强度值服从以下高斯分布:

$$I_{0n-A} \sim \mathcal{N}\left(\epsilon + i \frac{1-\epsilon}{N_1}, \frac{1-\epsilon}{N_1}\right)$$

设置 $A-Off$ 段上每帧的强度值服从以下高斯分布:

$$I_{A-Off} \sim \mathcal{N}\left(1 - (i - N_1) \frac{1-\epsilon}{N_2}, \frac{1-\epsilon}{N_2}\right)$$

设置峰值帧服从以下高斯分布:

$$I_A \sim \mathcal{N}\left(1, \min\left(\frac{1-\epsilon}{N_1}, \frac{1-\epsilon}{N_2}\right)\right)$$

式中, i 表示帧索引, $\mathcal{N}(\#, *)$ 表示均值为 $\#$ 方差为 $*$ 的高斯分布;

(3-2-4) 按照步骤(3-2-3)中强度值服从的高斯分布构造微表情帧序列的强度变化曲线。

5. 根据权利要求1所述的样本自适应的微表情放大方法, 其特征在于: 所述采用秩和检验的方式, 将对于微表情视频序列提取到的强度值按照所述强度变化曲线的趋势进行校正, 具体包括:

(3-3-1) 对强度变化曲线按照帧序采样后得到强度值序列 $S_1 = (q_1, \dots, q_N)$, 并获取强度提取模块提取到的强度值序列 $S_2 = (k_1, \dots, k_N)$, k_* 、 q_* 分别表示微表情帧序列第 $*$ 帧的强度值, N 表示微表情帧序列的长度; 并建立零假设: 认为强度值序列 S_1 和强度值序列 S_2 具有相同分布, 显著性水平 $\alpha = 0.05$;

(3-3-2) 将两个强度值序列 S_1 、 S_2 混合在一起后, 获取强度值序列 S_1 、 S_2 中的每个元素在混合序列中的顺序, 即秩 r_i^q 、 r_i^k :

$$r_i^q = \text{rank of } q_i \text{ among } (q_1, \dots, q_N, k_1, \dots, k_N)$$

$$r_i^k = \text{rank of } k_i \text{ among } (q_1, \dots, q_N, k_1, \dots, k_N)$$

(3-3-3) 计算强度值序列 S_1 中的每个元素的秩的和, 记为 T_1 ; 计算 S_2 中的每个元素的秩的和, 记为 T_2 ;

(3-3-4) 按照下式计算强度值序列 S_2 的检验统计量 Z , 以及该 Z 值在标准正态分布中所对应的概率 p :

$$Z = \frac{T_2 - \mu_T}{\sigma_T}$$

$$p = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}$$

$$\text{式中, } \mu_T = \frac{N(2N+1)}{2}, \sigma_T = \sqrt{\frac{N^2(2N+1)}{12}};$$

(3-3-5) 采用以下公式计算强度校正的损失,从而在训练过程中实现强度值序列 S_2 逐渐向所述强度变化曲线的趋势校正:

$$\mathcal{L}_{\text{Wrst}} = \sum_{k \in D} \max(0, \xi - (p_k - \alpha))$$

式中, $\mathcal{L}_{\text{Wrst}}$ 表示强度校正的损失, D 表示微表情数据库的微表情视频数, p_k 表示第 k 个微表情视频对应的概率 p , ξ 表示超参, 取值为 0.1。

6. 根据权利要求 1 所述的样本自适应的微表情放大方法, 其特征在于: 所述将校正后的强度值与从微表情帧序列中提取到的人脸特征整合具体为:

(3-4-1) 将校正后的每个强度值伸展为一个强度向量;

(3-4-2) 将微表情帧序列的每帧输入 Resnet+FC 线性层的网络中, 提取每帧的人脸特征向量;

(3-4-3) 将强度向量和对应的帧的面部特征向量做逐元素相乘, 得到表情强度增强后的增强人脸特征向量。

7. 根据权利要求 1 所述的样本自适应的微表情放大方法, 其特征在于: 步骤 (4) 中训练时采用的损失函数具体为:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{ID} + \lambda_2 \mathcal{L}_{\text{Wrst}} + \lambda_3 \mathcal{L}_C$$

其中

$$\mathcal{L}_C = -\frac{1}{K} \sum_{k=1}^K \sum_{c=1}^C y_{kc} \log(p_{kc})$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1$$

\mathcal{L} 为总损失, $\lambda_1, \lambda_2, \lambda_3$ 分别为三个损失函数对应的权重, \mathcal{L}_{ID} 表示强度提取的损失, $\mathcal{L}_{\text{Wrst}}$ 表示强度校正的损失, \mathcal{L}_C 表示微表情特征向量交叉熵损失, K 为训练样本的总个数, C 为微表情总类别数, y_{kc} 为类别为 c 的第 k 个样本对应的标签值, p_{kc} 为其属于类别 c 的概率。

8. 一种样本自适应的微表情放大装置, 包括处理器及存储在存储器上并可在处理器上运行的计算机程序, 其特征在于: 所述处理器执行所述程序时实现权利要求 1-7 中任意一项所述的方法。

样本自适应的微表情放大方法及装置

技术领域

[0001] 本发明涉及微表情识别技术,尤其涉及一种样本自适应的微表情放大方法及装置。

背景技术

[0002] 微表情(ME)是一种自发的、短暂的面部表情,通常由有意识地掩饰真实的情绪状态引起。与宏表情相比,微表情的运动具有局部性,并且强度很低,持续时间也很短(不到0.5秒),很难被发现和识别。这些特点使得准确地判定微表情的类别极为困难,但同时,微表情识别在健康诊断、国土安全、刑事审讯等诸多领域有着不可或缺的作用。这些潜在的应用吸引了越来越多的研究者进入这一领域,并提出了越来越多的方法来提高微表情识别的性能。

[0003] 在这些方法中,卷积神经网络(convolutional neural networks,CNN)被广泛应用于提取具有判别性的微表情特征。然而,微表情的低强度使得运动特征的特征变得困难。为了解决这一问题,许多方法采用运动放大技术对微表情强度进行放大,使其运动更加显著。一些传统的放大技术和深度学习滤波器,如欧拉运动放大(EMM)、全局拉格朗日运动放大(GLMM)、基于学习的运动放大(LMM),已经显示出增加微表情强度有助于提高识别的准确率。

[0004] 然而,由于不同被试呈现不同表情时面部肌肉的变化趋势存在差异,因此统一放大策略不能适合所有微表情样本。例如,对某一个样本设置的相同放大因子可能不足以提取鉴别特征,而对另一个对象设置的放大因子可能过多而引入噪声。因此,对于一个数据库中的所有微表情样本寻求统一的放大倍数是难以实现的。此外,微表情具有局部面部运动的特点,所以有些动作可能与识别特定的微表情类别无关,如眨眼、摇头等。这些限制促使寻求一种适合于每个特定微表情的放大方法,从而提高识别准确性。

发明内容

[0005] 发明目的:本发明针对现有技术存在的问题,提供一种可以根据每个微表情自适应放大的微表情放大方法及装置。

[0006] 技术方案:本发明所述的基于样本自适应的微表情放大方法包括:

[0007] (1) 获取微表情数据库,所述微表情数据库中包含有若干微表情视频和对应的微表情类别标签;

[0008] (2) 将微表情视频转换为微表情帧序列,并对序列长度进行统一;

[0009] (3) 构建微表情放大模型,所述微表情放大模型包括:

[0010] 强度提取模块,用于利用微表情帧序列之间的强度差异,通过对比学习的方式提取强度值;

[0011] 强度变化曲线构造模块,用于基于高斯分布构造微表情帧序列的强度变化曲线;

[0012] 强度校正模块,用于采用秩和检验的方式,将所述强度提取模块提取到的强度值

按照所述强度变化曲线的趋势进行校正；

[0013] 人脸增强模块,用于将校正后的强度值与从微表情帧序列中提取到的人脸特征整合,得到增强人脸特征向量；

[0014] 微表情特征整合模块,用于将属于一个微表情视频的所有帧的增强人脸特征向量使用LSTM整合,作为放大后的微表情特征向量；

[0015] (4)将统一长度的微表情帧序列和对应的微表情类别标签作为样本输入微表情放大模型进行训练,训练时采用的损失函数包括微表情特征向量交叉熵损失、强度提取的损失和强度校正的损失；

[0016] (5)将待放大的微表情视频进行长度统一后输入训练好的微表情放大模型,得到放大后的微表情特征向量。

[0017] 进一步的,步骤(2)中所述对序列长度进行统一具体包括：

[0018] 从每个微表情帧序列中按照预设间隔抽取N帧,形成长度都为N的微表情帧序列,N为正整数。

[0019] 进一步的,所述利用微表情帧序列之间的强度差异,通过对比学习的方式提取强度值,具体包括：

[0020] (3-1-1)对于统一后的每个微表情帧序列,将每帧x作为源样本,对x使用数据增强生成两个新的图像 \tilde{x}_1 和 \tilde{x}_2 ,作为x的正样本,从序列除x外的剩余帧中选择M张图片,记为 $x_1^-, x_m^-, \dots, x_M^-$,作为x的负样本,M为小于N的正整数,N表示微表情帧序列的长度；

[0021] (3-1-2)将每帧的正样本 \tilde{x}_1, \tilde{x}_2 和负样本 $x_1^-, x_m^-, \dots, x_M^-$ 依次输入Resnet网络+FC线性层的网络中,得到特征向量 \tilde{z}_1, \tilde{z}_2 和 $z_1^-, z_m^-, \dots, z_M^-$ ；

[0022] (3-1-3)将每帧的特征向量 $\tilde{z}_1, \tilde{z}_2, z_1^-, z_m^-, \dots, z_M^-$ 输入一线性层,得到对应的强度值；

[0023] (3-1-4)采用以下公式计算强度提取的损失,从而实现训练过程中强度信息提取的优化：

$$[0024] \quad \mathcal{L}_{ID} = \sum_{k \in D} \frac{-1}{N} \sum_{i=1}^N \left[\log \frac{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i})}{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}) + \sum_{m=1}^M \text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_m^{-k,i})} + \log \frac{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i})}{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}) + \sum_{m=1}^M \text{sim}(\tilde{z}_2^{k,i}, \tilde{z}_m^{-k,i})} \right]$$

[0025] 式中,i表示帧索引,D表示微表情数据库的微表情视频数,k表示微表情视频的索引, $\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}, \tilde{z}_m^{-k,i}$ 分别表示微表情数据库中第k个微表情视频长度统一后的第i帧的特征向量 $\tilde{z}_1, \tilde{z}_2, z_m^-$ 。

[0026] 进一步的,所述基于高斯分布构造微表情帧序列的强度变化曲线具体包括：

[0027] (3-2-1)设定微表情帧序列强度的变化范围在 $[\epsilon, 1]$,其中 $\epsilon > 0$ ；

[0028] (3-2-2)以峰值帧为界限将微表情帧序列分为两段,从起始帧到峰值帧记为 0_n-A 段,从峰值帧到结束帧记为 $A-0_{ff}$ 段,其中 0_n-A 段含有 N_1 帧, $A-0_{ff}$ 段含有 N_2 帧, $N_1+N_2+1=N$,N表示微表情帧序列的长度；

[0029] (3-2-3)设置 0_n-A 段上每帧的强度值服从以下高斯分布：

$$[0030] \quad I_{0_n-A} \sim \mathcal{N} \left(\epsilon + i \frac{1-\epsilon}{N_1}, \frac{1-\epsilon}{N_1} \right)$$

[0031] 设置A-Off段上每帧的强度值服从以下高斯分布：

$$[0032] \quad I_{A-Of} \sim \mathcal{N}\left(1 - (i - N_1) \frac{1 - \epsilon}{N_2}, \frac{1 - \epsilon}{N_2}\right)$$

[0033] 设置峰值帧服从以下高斯分布：

$$[0034] \quad I_A \sim \mathcal{N}\left(1, \min\left(\frac{1 - \epsilon}{N_1}, \frac{1 - \epsilon}{N_2}\right)\right)$$

[0035] 式中, i 表示帧索引, $\mathcal{N}(\#, *)$ 表示均值为#方差为*的高斯分布；

[0036] (3-2-4) 按照步骤 (3-2-3) 中强度值服从的高斯分布构造微表情帧序列的强度变化曲线。

[0037] 进一步的, 所述采用秩和检验的方式, 将对于微表情视频序列提取到的强度值按照所述强度变化曲线的趋势进行校正, 具体包括：

[0038] (3-3-1) 对强度变化曲线按照帧序采样后得到强度值序列 $S_1 = (q_1, \dots, q_N)$, 并获取强度提取模块提取到的强度值序列 $S_2 = (k_1, \dots, k_N)$, k_* 、 q_* 分别表示微表情帧序列第*帧的强度值, N 表示微表情帧序列的长度；并建立零假设：认为强度值序列 S_1 和强度值序列 S_2 具有相同分布, 显著性水平 $\alpha = 0.05$ ；

[0039] (3-3-2) 将两个强度值序列 S_1 、 S_2 混合在一起后, 获取强度值序列 S_1 、 S_2 中的每个元素在混合序列中的顺序, 即秩 r_i^q 、 r_i^k ：

$$[0040] \quad r_i^q = \text{rank of } q_i \text{ among } (q_1, \dots, q_N, k_1, \dots, k_N)$$

$$[0041] \quad r_i^k = \text{rank of } k_i \text{ among } (q_1, \dots, q_N, k_1, \dots, k_N)$$

[0042] (3-3-3) 计算强度值序列 S_1 中的每个元素的秩的和, 记为 T_1 ；计算 S_2 中的每个元素的秩的和, 记为 T_2 ；

[0043] (3-3-4) 按照下式计算强度值序列 S_2 的检验统计量 Z , 以及该 Z 值在标准正态分布中所对应的概率 p ：

$$[0044] \quad Z = \frac{T_2 - \mu_T}{\sigma_T}$$

$$[0045] \quad p = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}$$

$$[0046] \quad \text{式中, } \mu_T = \frac{N(2N+1)}{2}, \quad \sigma_T = \sqrt{\frac{N^2(2N+1)}{12}};$$

[0047] (3-3-5) 采用以下公式计算强度校正的损失, 从而在训练过程中实现强度值序列 S_2 逐渐向所述强度变化曲线的趋势校正：

$$[0048] \quad \mathcal{L}_{\text{Wrst}} = \sum_{k \in D} \max(0, \xi - (p_k - \alpha))$$

[0049] 式中, $\mathcal{L}_{\text{Wrst}}$ 表示强度校正的损失, D 表示微表情数据库的微表情视频数, p_k 表示第 k 个微表情视频对应的概率 p , ξ 表示超参, 取值为 0.1。

[0050] 进一步的, 所述将校正后的强度值与从微表情帧序列中提取到的人脸特征整合具体为：

[0051] (3-4-1) 将校正后的每个强度值伸展为一个强度向量；

[0052] (3-4-2) 将微表情帧序列的每帧输入Resnet+FC线性层的网络中,提取每帧的人脸特征向量;

[0053] (3-4-3) 将强度向量和对应的帧的面部特征向量做逐元素相乘,得到表情强度增强后的增强人脸特征向量。

[0054] 进一步的,步骤(4)中训练时采用的损失函数具体为:

$$[0055] \quad \mathcal{L} = \lambda_1 \mathcal{L}_{ID} + \lambda_2 \mathcal{L}_{Wrst} + \lambda_3 \mathcal{L}_C$$

[0056] 其中

$$[0057] \quad \mathcal{L}_C = -\frac{1}{K} \sum_{k=1}^K \sum_{c=1}^C y_{kc} \log(p_{kc})$$

$$[0058] \quad \lambda_1 + \lambda_2 + \lambda_3 = 1$$

[0059] \mathcal{L} 为总损失, $\lambda_1, \lambda_2, \lambda_3$ 分别为三个损失函数对应的权重, \mathcal{L}_{ID} 表示强度提取的损失, \mathcal{L}_{Wrst} 表示强度校正的损失, \mathcal{L}_C 表示微表情特征向量交叉熵损失,K为训练样本的总个数,C为微表情总类别数, y_{kc} 为类别为c的第k个样本对应的标签值, p_{kc} 为其属于类别c的概率。

[0060] 本发明所述的样本自适应的微表情放大装置,包括处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现上述方法。

[0061] 有益效果:本发明与现有技术相比,其显著优点是:本发明摒弃了现有的微表情识别技术中普遍采用的在图像空间上放大表情强度的方式,提供了一种新的放大思路。本发明的目的是使得放大信息能够自适应的根据不同的样本进行调节,且随着网络的优化而动态改变,具体地,本发明通过对比学习提取强度信息,随后通过秩和检验的方式校正强度信息的变化方式,从而实现在特征空间上的强度信息表征,实验结果证明,本发明可以提高识别准确率,相较于现有的放大方式,减少了人工干预且增加了可调整性。

附图说明

[0062] 图1是本发明提供的样本自适应的微表情放大方法的一个实施例的流程示意图;

[0063] 图2是本发明设计的整个过程的框架图;

[0064] 图3是强度信息提取的示意图;

[0065] 图4是采用秩和检验进行强度信息校正的示意图。

具体实施方式

[0066] 本实施例提供一种基于样本自适应的微表情放大方法,如图1和图2所示,包括如下步骤:

[0067] (1) 获取微表情数据库,所述微表情数据库中包含有若干微表情视频和对应的微表情类别标签。

[0068] (2) 将微表情视频转换为微表情帧序列,并对序列长度进行统一。

[0069] 由于每个微表情视频样本在获取的时候有不同的帧数,为了方便网络识别,需要将所有视频的长度统一,参考大多数视频预处理的方式,统一从每个微表情视频样本中选取 $N=16$ 帧,选取的方式是根据视频的长度以一定的间隔从样本中抽取帧,例如,如果原视频有 N_0 帧,则以 $N_0/16$ 的间隔选出帧,保证视频的每个阶段都有帧被选到。

[0070] (3) 构建微表情放大模型,所述微表情放大模型包括强度提取模块、强度变化曲线构造模块、强度校正模块、人脸增强模块、微表情特征整合模块。

[0071] 强度提取模块用于利用微表情帧序列之间的强度差异,通过对比学习的方式提取强度值。如图3所示,具体包括:

[0072] (3-1-1) 对于统一后的每个微表情帧序列,将每帧 x 作为源样本,对 x 使用数据增强生成两个新的图像 \tilde{x}_1 和 \tilde{x}_2 ,作为 x 的正样本,从序列除 x 外的剩余帧中选择 M 张图片,记为 $x_1^-, x_m^-, \dots, x_M^-$,作为 x 的负样本, M 为小于 N 的正整数, N 表示微表情帧序列的长度;

[0073] (3-1-2) 将每帧的正样本 \tilde{x}_1, \tilde{x}_2 和负样本 $x_1^-, x_m^-, \dots, x_M^-$ 依次输入Resnet网络+FC线性层的网络中,得到特征向量 \tilde{z}_1, \tilde{z}_2 和 $z_1^-, z_m^-, \dots, z_M^-$;

[0074] (3-1-3) 将每帧的特征向量 $\tilde{z}_1, \tilde{z}_2, z_1^-, z_m^-, \dots, z_M^-$ 输入一线性层,得到对应的强度值;

[0075] (3-1-4) 采用以下公式计算强度提取的损失,从而实现训练过程中强度信息提取的优化:

$$[0076] \quad \mathcal{L}_{ID} = \sum_{k \in D} \frac{-1}{N} \sum_{i=1}^N \left[\log \frac{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i})}{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}) + \sum_{m=1}^M \text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_m^{-k,i})} + \log \frac{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i})}{\text{sim}(\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}) + \sum_{m=1}^M \text{sim}(\tilde{z}_2^{k,i}, \tilde{z}_m^{-k,i})} \right]$$

[0077] 式中, i 表示帧索引, D 表示微表情数据库的微表情视频数, k 表示微表情视频的索引, $\tilde{z}_1^{k,i}, \tilde{z}_2^{k,i}, \tilde{z}_m^{-k,i}$ 分别表示微表情数据库中第 k 个微表情视频长度统一后的第 i 帧的特征向量 $\tilde{z}_1, \tilde{z}_2, z_m^-$ 。

[0078] 强度变化曲线构造模块用于基于高斯分布构造微表情帧序列的强度变化曲线。在原始的微表情视频序列中,强度通常按照先增后减的趋势,这一点已被诸多微表情有关的文献证实,根据这一趋势为每个微表情序列构造变化曲线,曲线的具体值并不重要,重要的是模拟先增后减的趋势。因此,构建的方法具体包括:

[0079] (3-2-1) 设定微表情帧序列强度的变化范围在 $[\epsilon, 1]$,其中 $\epsilon > 0$;

[0080] (3-2-2) 以峰值帧为界限将微表情帧序列分为两段,从起始帧到峰值帧记为 $0n-A$ 段,从峰值帧到结束帧记为 $A-Off$ 段,其中 $0n-A$ 段含有 N_1 帧, $A-Off$ 段含有 N_2 帧, $N_1+N_2+1=N$, N 表示微表情帧序列的长度;

[0081] (3-2-3) 设置 $0n-A$ 段上每帧的强度值服从以下高斯分布:

$$[0082] \quad I_{0n-A} \sim \mathcal{N}\left(\epsilon + i \frac{1-\epsilon}{N_1}, \frac{1-\epsilon}{N_1}\right)$$

[0083] 设置 $A-Off$ 段上每帧的强度值服从以下高斯分布:

$$[0084] \quad I_{A-Off} \sim \mathcal{N}\left(1 - (i - N_1) \frac{1-\epsilon}{N_2}, \frac{1-\epsilon}{N_2}\right)$$

[0085] 设置峰值帧服从以下高斯分布:

$$[0086] \quad I_A \sim \mathcal{N}\left(1, \min\left(\frac{1-\epsilon}{N_1}, \frac{1-\epsilon}{N_2}\right)\right)$$

[0087] 式中, i 表示帧索引, $\mathcal{N}(\#,*)$ 表示均值为#方差为*的高斯分布;

[0088] (3-2-4) 按照步骤(3-2-3)中强度值服从的高斯分布构造微表情帧序列的强度变

化曲线。

[0089] 强度校正模块用于采用秩和检验的方式,将所述强度提取模块提取到的强度值按照所述强度变化曲线的趋势进行校正。如图4所示,具体包括:

[0090] (3-3-1)对强度变化曲线按照帧序采样后得到强度值序列 $S_1 = (q_1, \dots, q_N)$,并获取强度提取模块提取到的强度值序列 $S_2 = (k_1, \dots, k_N)$, k_* 、 q_* 分别表示微表情帧序列第*帧的强度值, N 表示微表情帧序列的长度;并建立零假设:认为强度值序列 S_1 和强度值序列 S_2 具有相同分布,显著性水平 $\alpha = 0.05$;

[0091] (3-3-2)将两个强度值序列 S_1 、 S_2 混合在一起后,获取强度值序列 S_1 、 S_2 中的每个元素在混合序列中的顺序,即秩 r_i^q 、 r_i^k :

[0092] $r_i^q = \text{rank of } q_i \text{ among } (q_1, \dots, q_N, k_1, \dots, k_N)$

[0093] $r_i^k = \text{rank of } k_i \text{ among } (q_1, \dots, q_N, k_1, \dots, k_N)$

[0094] (3-3-3)计算强度值序列 S_1 中的每个元素的秩的和,记为 T_1 ;计算 S_2 中的每个元素的秩的和,记为 T_2 ;

[0095] (3-3-4)按照下式计算强度值序列 S_2 的检验统计量 Z ,以及该 Z 值在标准正态分布中所对应的概率 p :

[0096]
$$Z = \frac{T_2 - \mu_T}{\sigma_T}$$

[0097]
$$p = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}$$

[0098] 式中, $\mu_T = \frac{N(2N+1)}{2}$ 、 $\sigma_T = \sqrt{\frac{N^2(2N+1)}{12}}$;

[0099] 其中, Z 值的计算主要是利用在序列中的样本个数 $N > 10$ 的情况下,认为总秩和 T 服从如下的高斯分布的原理:

[0100]
$$T \sim \mathcal{N}(\mu_T, \sigma_T)$$

$$= \mathcal{N}\left(\frac{N(2N+1)}{2}, \sqrt{\frac{N^2(2N+1)}{12}}\right)$$

[0101] (3-3-5)采用以下公式计算强度校正的损失,从而在训练过程中实现强度值序列 S_2 逐渐向所述强度变化曲线的趋势校正:

[0102]
$$\mathcal{L}_{\text{Wrst}} = \sum_{k \in D} \max(0, \xi - (p_k - \alpha))$$

[0103] 式中, $\mathcal{L}_{\text{Wrst}}$ 表示强度校正的损失, D 表示微表情数据库的微表情视频数, p_k 表示第 k 个微表情视频对应的概率 p , ξ 表示超参,取值为0.1。

[0104] 人脸增强模块用于将校正后的强度值与从微表情帧序列中提取到的人脸特征整合,得到增强人脸特征向量。

[0105] 微表情特征整合模块用于将属于一个微表情视频的所有帧的增强人脸特征向量使用LSTM整合,作为放大后的微表情特征向量。具体包括:

[0106] (3-4-1)将校正后的每个强度值伸展为一个强度向量;

[0107] (3-4-2)将微表情帧序列的每帧输入Resnet+FC线性层的网络中,提取每帧的人脸

特征向量；

[0108] (3-4-3) 将强度向量和对应的帧的面部特征向量做逐元素相乘，得到表情强度增强后的增强人脸特征向量。

[0109] (4) 将统一长度的微表情帧序列和对应的微表情类别标签作为样本输入微表情放大模型进行训练，训练时采用的损失函数包括微表情特征向量交叉熵损失、强度提取的损失和强度校正的损失。

[0110] 采用的损失函数具体为：

$$[0111] \quad \mathcal{L} = \lambda_1 \mathcal{L}_{ID} + \lambda_2 \mathcal{L}_{Wrst} + \lambda_3 \mathcal{L}_C$$

[0112] 其中

$$[0113] \quad \mathcal{L}_C = -\frac{1}{K} \sum_{k=1}^K \sum_{c=1}^C y_{kc} \log(p_{kc})$$

$$[0114] \quad \lambda_1 + \lambda_2 + \lambda_3 = 1$$

[0115] \mathcal{L} 为总损失， $\lambda_1, \lambda_2, \lambda_3$ 分别为三个损失函数对应的权重， \mathcal{L}_{ID} 表示强度提取的损失， \mathcal{L}_{Wrst} 表示强度校正的损失， \mathcal{L}_C 表示微表情特征向量交叉熵损失， K 为训练样本的总个数， C 为微表情总类别数， y_{kc} 为类别为 c 的第 k 个样本对应的标签值， p_{kc} 为其属于类别 c 的概率。

[0116] (5) 将待放大的微表情视频进行长度统一后输入训练好的微表情放大模型，得到放大后的微表情特征向量。

[0117] 本实施例还提供了一种样本自适应的微表情识别装置，包括处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述程序时实现上述方法。

[0118] 为验证本发明的有效性，在CASME2微表情数据库、SAMM微表情数据库和SMIC-HS数据库上做了微表情识别，验证结果及与其他最新方法的比较如表1所示：

[0119] 表1

[0120]

方法	年份	CASME2	SAMM	SMIC-HS
LBP-SIP	2014	45.36	36.76	42.12
MagGA	2018	63.30	N/A	N/A
DSSN	2019	70.78	57.35	63.41
TSCNN-I	2020	74.05	63.53	72.74
LBPAccP ^{u2}	2021	69.03	N/A	76.59
AU-GCN	2021	74.27	74.26	N/A
本方法	2022	78.05	78.68	79.27

[0121] 表1中，N/A表示没有相关记录。

[0122] 对CASME2数据库的表情做如下处理：只保留样本数目在10个以上的类别，即happy、regression、disgust、fear和surprise。同样地，对SAMM数据库，只保留样本数目在10个以上的类别，分别是happiness、angry、disgust、fear和surprise。SMIC数据库的类别有positive、negative、surprise，所以在该数据库上执行三分类的任务。

[0123] 实验结果表明，基于本发明提出的微表情识别方法具有较好的微表情识别性能。相较于传统的微表情放大方式，本发明更能够适应不同微表情样本之间的差异，执行起来也更方便。

[0124] 以上所揭露的仅为本发明一种较佳实施例而已,不能以此来限定本发明之权利范围,因此依本发明权利要求所作的等同变化,仍属本发明所涵盖的范围。

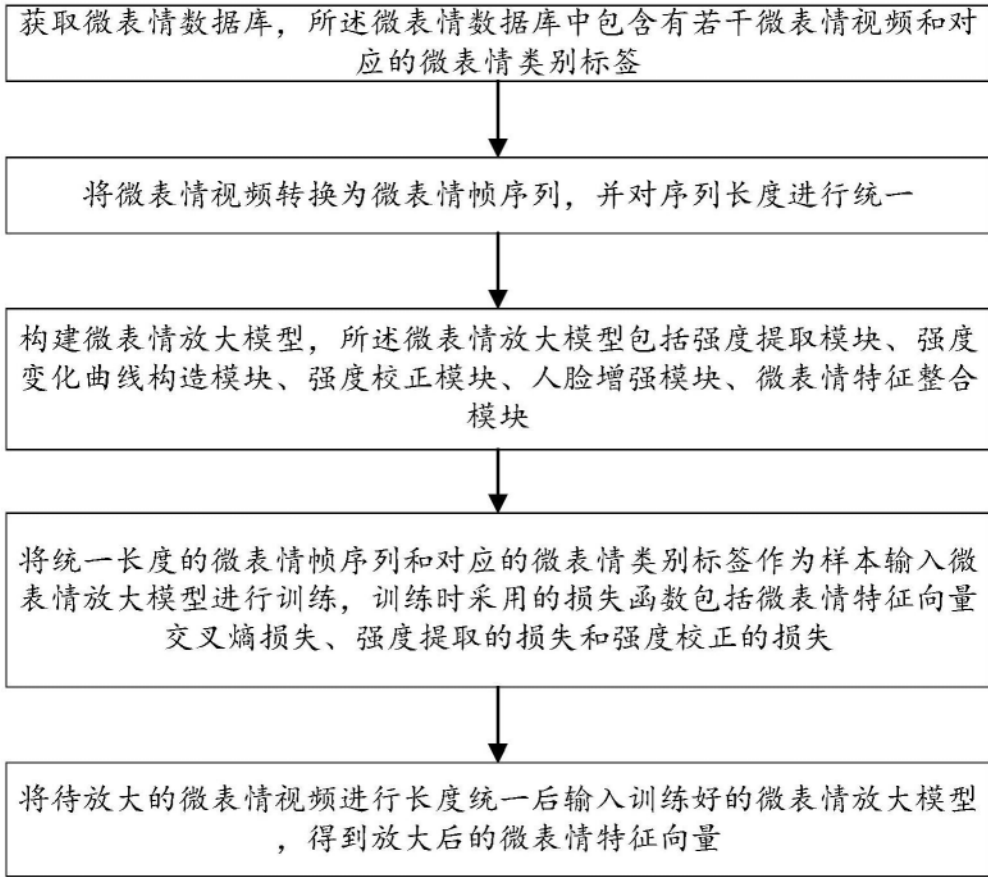


图1

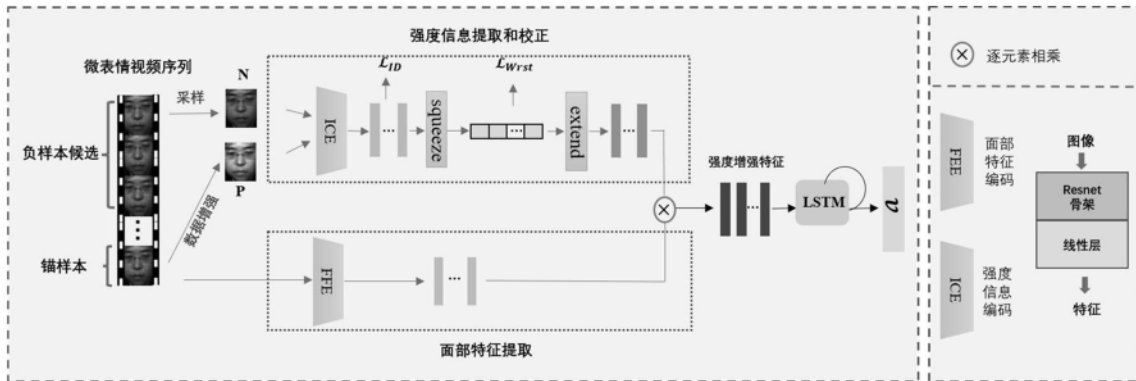


图2

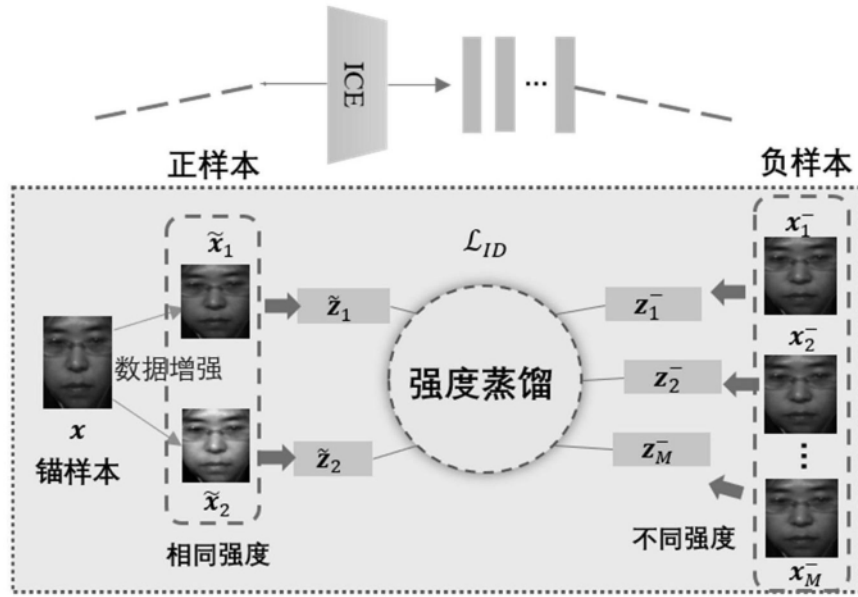


图3

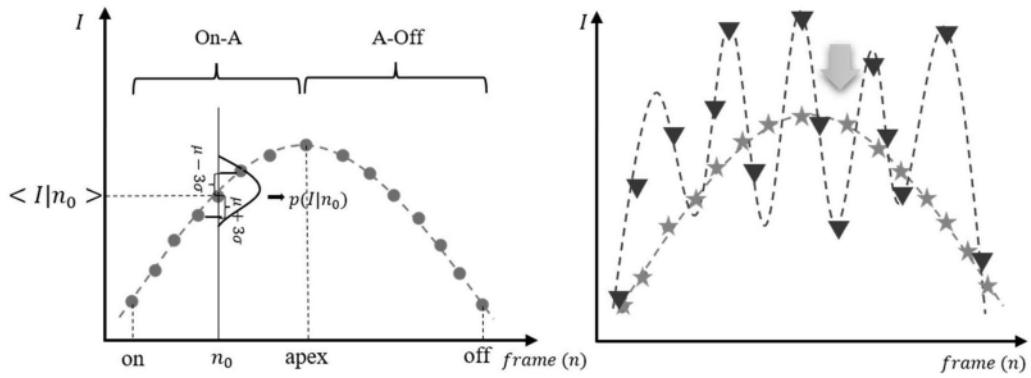


图4