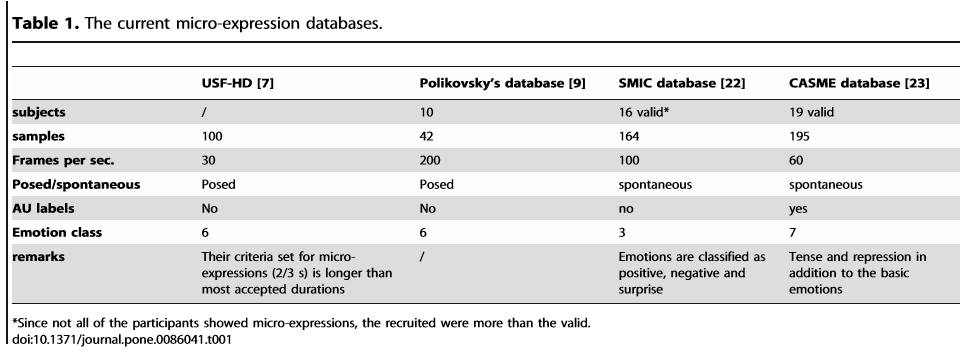
# 摘要

文章回顾了以前开发的微表情数据库，并建立了改进的微表情数据库（CASMEII），具有更高的时间分辨率（200fps）和空间分辨率（面部区域约2806340像素）。我们在良好控制的实验室环境中和适当的照明条件（例如消除灯光闪烁）下激发了参与者的面部表情。在近3000个面部动作中，为数据库选择了247个微表情，并标记了动作单位（AU）和情绪。对于基线（基准）评估，分别采用LBP-TOP和SVM进行特征提取和分类，并采用留一法则交叉验证方法。

# 引言

自动面部表情识别技术正在蓬勃发展。研究人员开发了许多算法，对六种基本面部表情（愤怒，厌恶，恐惧，幸福，悲伤和惊奇）的准确性已达到90％以上。相反，自动微表情识别的研究最近才开始。Shreve[7]等人将脸部划分为各个子区域（嘴，脸颊，前额和眼睛），并计算每个子区域的面部应变，然后分析每个子区域中的应变模式，以检测视频剪辑中的微表情。Pfister[8]等提出了一个使用时间插值模型和多核学习来识别微表情的框架。Polikovsky[9]等使用3D梯度描述符进行微表情识别。吴等[10]通过使用Gabor功能并使用GentleSVM作为分类器，开发了一种自动微表情识别系统。Wang[11]等将灰度微表情视频片段视为三阶张量，并利用判别张量子空间分析（DTSA）和极限学习机（ELM）来识别微表情。RuizHernandez和Pietikaéinen[12]提出使用二次局部高斯射流的重新参数化来编码局部二进位模式（LBP），以生成更鲁棒和可靠的直方图用于微表情表示。

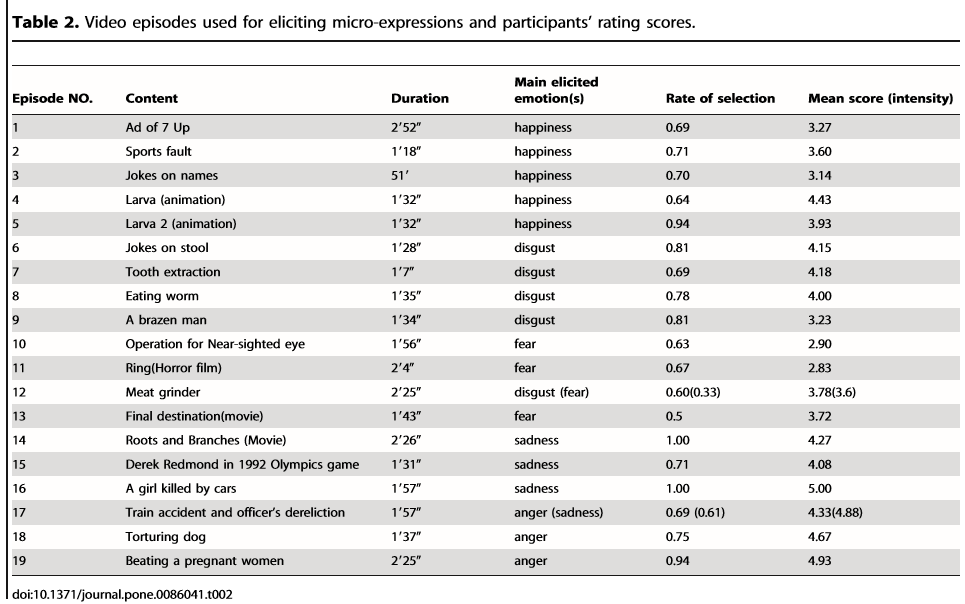
完善的微表情数据库很少，这阻碍了微表情识别研究的发展。到目前为止，只有四个微表情数据库，其中只有两个包含自发的微表情，如表1所示。

开发新数据库的原因：

# 方法

## 实验材料

使用类似于在CASME中引起微表情的视频剪辑片段，但是删除了几个无效的情节，并添加了几个新的情节。20名参与者通过从列表中选择一个或两个情感关键词并以7点李克特量表(Likertscale)（最低为0，最高为6）对其强度进行评分来对视频片段进行了评分，如果对与某个视频片段，志愿者中有三分之一或更多的人选择了某种情感关键词（例如幸福感），则该情感关键词将被视为该视频片段的主要情感。在相同的视频片段中，主观感受存在个体差异同时，一个视频片段可能引起几种情绪，这在视频片段＃12和＃17中得到明显体现。表2列出了所选视频片段的详细信息及其相应的评分分数。



## 实验程序

### 记录

18位参与者在观看视频剪辑时被要求保持中和的脸，而17位参与者仅在意识到存在面部表情时才试图抑制面部运动，如上引言中所述，此设计用于引发两种类型的微表情[1]。要求志愿者在屏幕前观看视频剪辑，并避免任何身体动作，研究者通过另一台显示器在线监控志愿者的脸部。

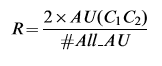
### 样品选择和类别标签

两个编码器参与了微表情的分析。我们按照以下步骤处理原始视频录像，以选择微表情：

**步骤1**：去除无关的面部动作。录音中有大量的面部动作，但很多与面部动作无关，例如头部动作、吞咽唾液时按压嘴唇或其他习惯性动作（例如吹鼻子），所有没有感情的运动都被排除在外。

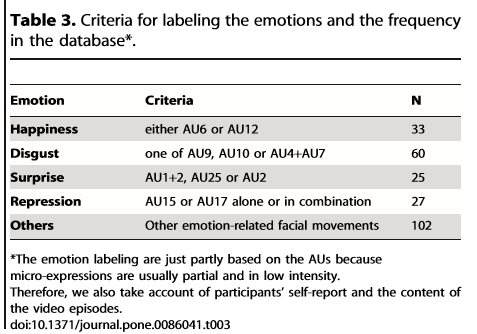
**步骤2**：通过限制要分析的样本来预先选择微表情（分析的）候选对象。编码器浏览并找到每个情绪动作片段的开始和偏移帧，只有持续时间少于一秒钟的那些被保存为微表情候选，以进行进一步的编码，过于细微而无法精确编码的表情也被排除在外。

**步骤3**：将候选微表情片段转换为帧序列，以利于精确标记起始和偏移帧。通过采用逐帧方法，确定每个序列的开始和偏移的精确帧数，只有满足选择标准的那些（总持续时间小于500ms或开始持续时间小于250ms）才被选最终的微表情样品。选好微表情样本后，标记AUs(动作单元)，并由两个编码器为每个微表情提供情感标签。可靠性为0.846，其计算公式为：



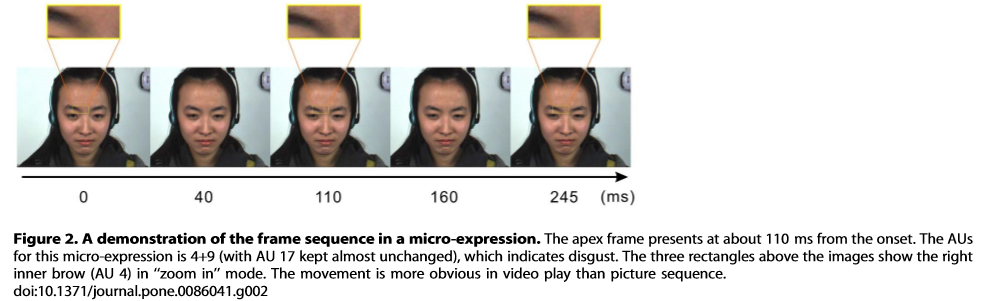
其中*#AU（）*是编码器1和编码器2达成共识（都同意是某个AU?）的AU数，而*#All\_AU*是两个编码器在微表情式中计分的AU总数。之后，两个编码器讨论并仲裁了分歧。

在CASMEII数据库中，我们基于AU、参与者的自我报告以及视频片段的内容来标记（分类）微表情。我们提供了五个主要类别（如表3所示）。



### 数据库简要介绍

开发的数据库CASMEII包含来自26位参与者的247个微表情样品，它们是从近3,000种引发的面部动作中选择出来的。这些样本使用开始帧和偏移帧进行编码，标记了动作单位（AU）和情感标签，并提供了五种主要类别，如表3所示。表3还显示了用于标记情感类别的标准，图2显示了其中一种微表情片段的示例。



在这个数据库中，删除了非常微妙的微表情，因为它们几乎无法编码起始帧和偏移帧（由于几乎察觉不到的变化，难以检测转折点）。CASMEII数据库具有以下特征：

* 这些微表情样品都是自发的和动态的。在每次微表情之前和之后都会保留基线（通常是中性的）帧，从而可以评估不同的检测算法。
* 具有较高的时间分辨率（200fps）和相对较高的面部分辨率（280X340像素）。
* 微表情标记基于FACS研究人员指南[24]和Yan等人的发现[3]，标记标准与普通6个类别的普通面部表情不同。
* 录音具有适当的照明，不会闪烁光线，并且减少面部的高光区域。
* 在实验室情况下，某些类型的面部表情难以激发，因此不同类别的样本分布不均，例如，有60个令人厌恶的样本，但只有7个悲伤样本。在CASMEII中，我们提供5类微表情式。

# 数据库评估

## 预处理

在此数据集上进行任何正式实验之前，对已标记的原始样品片段进行了三个预处理步骤。令S=[|*s*∈S,*i*=1,...,n]为微表情片段的集合。第i个样本=[|*f*∈,*j*=1,…,]，是序列的帧号。

**首先**，选择具有中性表情的正面面部图像M作为模型面部（基准）。使用活动形状模型(ASM)(ActiveShapeModel)[25]可以检测到模型面部（基准）ψ(M)的68个面部标记点。**其次**，使用局部加权均值（LWM）（LocalWeightedMean）[26]变换将每个微表情片段的第一帧标准化为模型面部（基准），变换矩阵T为：



其中ψ(,1)是微表情样本的第一帧的68个面部标记点的坐标。然后，使用相同的矩阵对的所有帧进行归一化。我们仅在第一帧而不是在所有帧上都检测ASM面部标记点有两个原因。第一个原因是，由于微表情的持续时间非常短，因此可以忽略该持续时间内严格的(?)头部的移动；第二个原因是ASM检测到的面部标记点可能不够准确，如果将其应用于一系列帧，即使面部完全不移动，同一点的位置也可能存在较大偏差。归一化图像f’被计算为原始图像的2D变换：

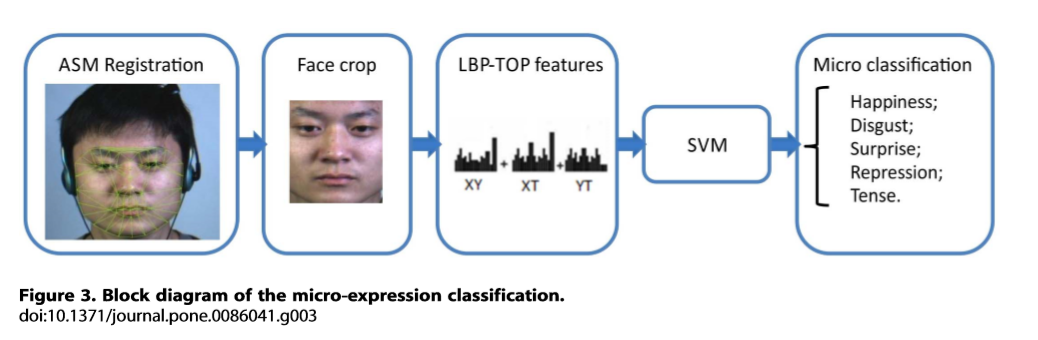


是归一化微表情序列的第j帧。

第三，定位每个归一化微表情序列的第一帧的眼睛坐标，然后使用由眼睛位置确定的矩形来裁剪出的每个帧的面部。

## 实验

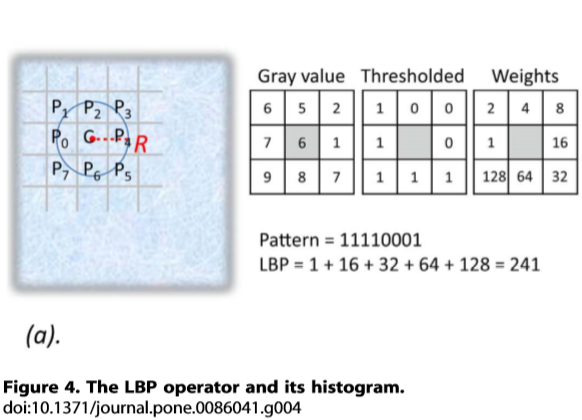
我们对通过上一节中描述的步骤做了预处理的样品进行了微表情识别实验。提取LBP-TOP特征以从时空的角度描述微表情，并使用SVM作为分类器，为我们的数据库的未来评估提供基线标准（图3），详细说明如下。



## 方法

### LocalBinaryPatternsfromThreeOrthogonalPlanes(LBP-TOP)(来自三个正交平面的局部二值模式)

事实证明，传统的LBP方法可有效描述静态图像的2D纹理[27]。考虑到图像中的一个像素C，LBP算子通过比较和阈值化其相邻像素的灰度值来描述其局部纹理图案。如图4（a）所示。



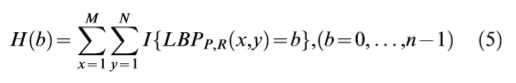
对于中心像素C及其P个相邻的半径R以内的像素，LBP值计算如下：



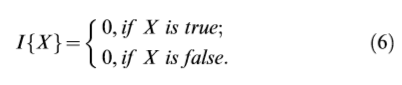
其中（）是中心像素C的坐标，是其灰度值，（p=0,...,P-1）是其在半径R内的第p个相邻像素的灰度值。是与相邻像素的位置相对应的权重，该权重用于将二进制模式字符串转换为此模式的十进制指数。函数s(x)是一个符号函数，定义为：



给定大小为MxN的输入图像，可以使用上述两个公式来计算每个像素的LBP值b。然后，我们可以通过计算所有LBP值的直方图分布来形成特征向量以表示输入图像：

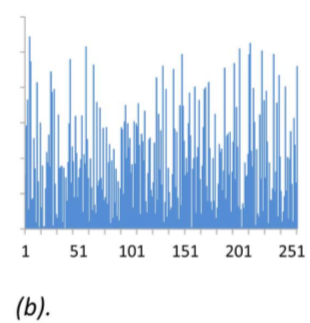


其中n表示感兴趣的模式总数，以及：



（应该是X为true时I{X}=1吧）

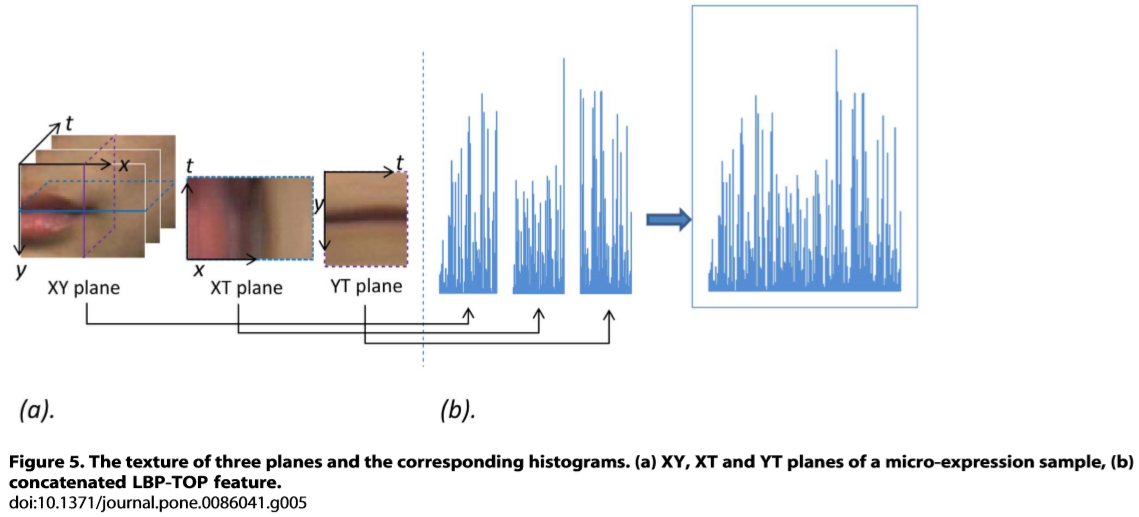
LBP直方图的一个示例在图4（b）中说明，有关LBP的更多详细信息请参阅[27]。



作为基本LBP概念的扩展，Zhao等人[28]提出了三个正交平面上的局部二值模式（LBP-TOP）用于时空域中的动态纹理分析。

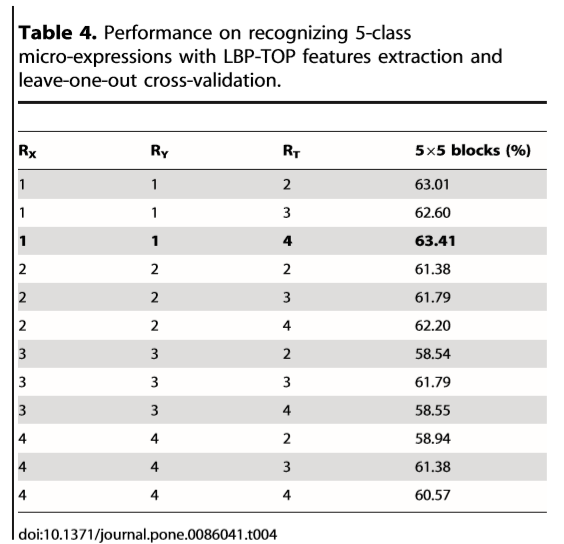
给定时间长度为T的视频序列，通常可以将其视为沿时间轴T的XY平面堆栈，但也可以将其视为在Y轴上的XT平面堆栈或在Y轴上的YT平面堆栈X。XT和YT平面提供有关时空转换的信息。以微表情为例，XT和YT平面包含有关像素的行或列的灰度值沿时间维度如何变化的信息。如果输入数据是嘴角，则它们将嘴部运动描述为XT或YT平面中的动态纹理（请参见图5（a））。

LBP-TOP的基本思想类似于LBP，LBP-TOP的基本思想是使用索引编码通过将其灰度值与其邻居的灰度值进行比较（并设定阈值/阈值化）来表示每个像素的局部模式。为了包括来自3D时空域的信息，分别从每个像素的XY，XT和YT平面提取LBP编码。然后，通过使用等式（3）和（4）形成每个平面的一个直方图，然后将其连接为单个直方图作为最终的LBP-TOP特征向量。请注意，我们可以为X，Y和T轴的半径设置不同的值，例如，和，因此最终可以进行椭圆采样而不是圆采样。图5（b）演示了提取LBP-TOP特征直方图的过程。



## 结果

X和Y轴的半径从1到4不等。T的半径从2到4不等（我们不考虑T=1，因为样本的采样率为200fps，其中两个相邻帧的变化很小），XY，XT和YT平面中相邻点的数量P设置为4，SVM被用作分类器。在此基线评估中，我们选择了5种微表情类别（快乐，惊讶，厌恶，压抑和其他）进行训练和测试，考虑到样本在五个选定类别中的分布不均，应用了留一法则交叉验证，在565个块中提取了LBP-TOP特征。表4中显示了性能，最佳性能为63.41％，其中XY，YT和XT平面的半径分别为1、1和4。



# 讨论与局限

关于激发微表情和该数据库，仍然存在一些限制：

**首先**，这些自发微表情达的标记不是很令人满意。引发微表情的材料是视频情节，它是复杂的刺激，对不同的人可能具有不同的含义，例如，咀嚼蠕虫的场景并不总是令人作呕，实际上有些志愿者认为其“有趣”或“有趣”。仅根据AU（特别是尚未定义的AU组合）对微表情进行分类需要进一步的深入研究。要求参与者控制和抑制其面部表情，这通常会使自发的微表情显得不完全（程度？），有时只有一个AU（另见Porter，2008；Yan等，2013）。即使我们考虑了AU组合、视频片段的内容和参与者的报告，这些微表情的情感标签仍然是有争议的。

**其次**，此数据库中的微表情是在一种特定的实验室情况下引起的，可能不包括在其他情况下引起的微表情。人们在警察审讯或向老板撒谎时出现的微表情可能与我们在实验室情况下观察到的微表情不同。到目前为止，很少有研究调查在不同情况下引起的微表情，未来我们应该进行进一步的研究，以探索不同情况下微表情的变化。

**第三**，目前尚不清楚哪种方法最适合对微表情式进行分类。微表情具有以下特征，这些特征与普通的面部表情在以下方面有所不同：（1）低强度的面部动作；（2）部分的面部表情（片段）。因此，先前的适用于对普通面部表情进行分类的方法可能不适用于微表情。我们提供了该数据库和基线评估，因此将来的研究可能会与我们的初步结果进行比较。

**第四**，参与者年龄有限。参加这个启发性范例实验的是年轻人，其中大多数是大学生。如果数据库包含不同年龄模型的样本，那将更好。