



基于 HMM 的面部表情图像序列的 分析与识别¹⁾

金辉¹ 高文^{1,2}

¹(哈尔滨工业大学 321 信箱计算机科学与工程系 618 实验室 哈尔滨 150001)

²(中国科学院计算技术研究所 北京 100080)

(E-mail: jhui@vilab.hit.edu.cn wgao@cti.com.cn)

关键词 动态表情, 表情建模与编码, 表情的特征流, 表情的分析与识别

中图分类号 TP391.4

ANALYSIS AND RECOGNITION OF FACIAL EXPRESSION IMAGE SEQUENCES BASED ON HMM

JIN Hui¹ GAO Wen^{1,2}

¹(P. O. BOX 321, Department of Computer Science and Engineering, Harbin Institute of Technology, Harbin 150001)

²(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

(E-mail: jhui@vilab.hit.edu.cn, wgao@cti.com.cn)

Key words Dynamic expression, modeling and coding of the facial expression, eigen flow of the expression, analysis and recognition of expression

1 引言

人们在实际生活中,依靠脸、手等体势信息的交流是一件最简单便利的事情,因而在计算机科学中,面部表情的识别对促进计算机视觉系统,建模和数据库的发展都有重要作用,是计算机情感计算的重要研究内容;在基础科学研究方面的语言学中和在行为学方面,以及在商业应用方面,如在可视电话,电视会议上都有重要的应用价值.

面部表情的研究是随着情绪理论而提出的.面部表情、声调表情和身体姿态三方面构成了情绪表现.而情绪表现、情绪体验和情绪生理这三种因素又组成了情绪的心理.因而表情是复杂的情绪心理的一个由脸部运动而产生的外在表现.国内,文献[1]曾把人脸识别方面特征脸(Eigenface)的思想应用到了表情识别中,但没有应用时间信息和运动信息;国际上,对人脸面部表情的识别及其相关方向的研究主要有美国的 MIT 的多媒体实验室

1) 国家自然科学基金(69789301)、国家“863”高技术研究发展计划项目(863-ZT03-01-2)和中国科学院“百人计划”基金资助

收稿日期 2000-08-04 收修改稿日期 2001-03-20

的感知计算组、马里兰大学的计算机视觉实验室,日本的城蹊大学和大阪大学等等. Yacoob 和 Davis 在八方向上检测运动,使用简化的 FACS 规则识别六种表情,对表情的分析是局部化的信息;Mark Rosenblum 和 Yaser Yacoob 等人^[2]用 RBFN 结构,学习脸部特征与人类情绪之间的相关性;Irfan A Essa 等^[3]把提取的新运动单元命名为“FACS+”,产生了一个随时间变化的脸部形状的空间模板和一个独立的肌肉运动群参数化表征;另外,James J Lien^[4]等应用三种方法:特征点跟踪,流跟踪和边缘检测对表情进行分析. 到目前为止还没有看到相关文献用特征流与 HMM 相结合来对表情图像序列作分析与识别.

2 动态表情的建模与编码

艾克曼用表情分类的方法定义了六种最基本表情:惊奇、恐惧、厌恶、愤怒、高兴、悲伤和 33 种表情倾向. 本文采用这种最基本的六种分类方法进行识别. Ekman 和 Friesen^[5]提出了目前最广泛被采用的、视觉可区分的人脸运动编码系统 FACS(Facial Action Coding System):是人脸上所有导致面部运动的运动单元的枚举. FACS 有两个主要弱点:1)运动单元是纯粹的局部化的空间模板;2)没有时间描述信息,不能精确描绘真实的脸部运动. FACS 中数据的分析只是单纯的描绘运动单元 AU(Action Units),或者把 FACS 通过字典规则转化成情绪. 本文在此基础上提出了 FACS'(转换 FACS)的表情编码,把运动单元的运动转化成基于物理结构和肌肉模型的运动特征向量序列来对眼部和嘴部分别进行表情编码,相应的运动解释基于 FACS 的规则,同时克服了 FACS 的弱点.

表情识别的一个难点,就是建立表情模型. 人脸是一个柔性体而不是一个刚体,很难把脸部的运动与表情联系起来. 本文根据表情图像序列的特点,建立了动态的表情模型,是一个完整的动态表情过程的描述,如图 1 所示.

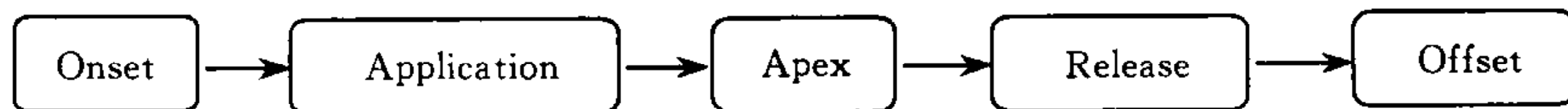


图 1 表情动态模型

3 表情特征区域的定位及标准化

人脸面部感知系统的前提条件是已知人脸图像. 人脸图像定位以后,根据面部结构物理-肌肉模型的先验信息,可以确定出各特征部件的大致区域,然后进一步提取表情的特征区域,它们是由表情变化比较显著的特征部件的相关肌肉定义的:眼睛区域,嘴部区域. 对同一组图像序列,特征区域的定位只在第一帧计算得到,其余帧与第一帧相同,并在标准化时使用了旋转与平移操作,以防止头部的刚体运动对面面部表情运动的影响;对不同组图像序列,各特征区域必须根据特征部件的位置与比例信息进行标准化. 在眼睛区域中,最具明显特征的是虹膜,因此很容易检测得到. 进而可得到眉毛特征,根据眼睛及眉毛的比例和位置信息,可以重新校正眼睛特征区域. 嘴部区域也是先从大致确定的粗定位区域中根据灰度积分特征找到嘴部,然后校正嘴部特征区域. 这样经过校正的特征区域,就保证了其中部件位置与比例的标准化. 然后对其大小做归一化处理^[6].

4 基于光流的表情运动的分析

表情特征区域的运动场是在特征区域基础上计算的. 关于运动向量即光流的估计, 主要有三种方法. 传统的 Horn 与 Schunck^[7]提出的基于梯度的方法, 比较适合于皮肤的变形计算, 而且计算量比较简单, 它只是逐点估计位置的瞬时速度场, 本文采用这种方法. 在光流计算中, 先对图像序列进行时间空间上的平滑, 再对光流场进行平滑, 以使其能精确描绘特征的运动. 如图 2, 分别是眼部和嘴部特征区域的光流场. 用光流特征作为特征向量, 对光照均匀程度的鲁棒性较强, 只要光照在同一组序列中强度不变, 提取的特征就是稳定的.

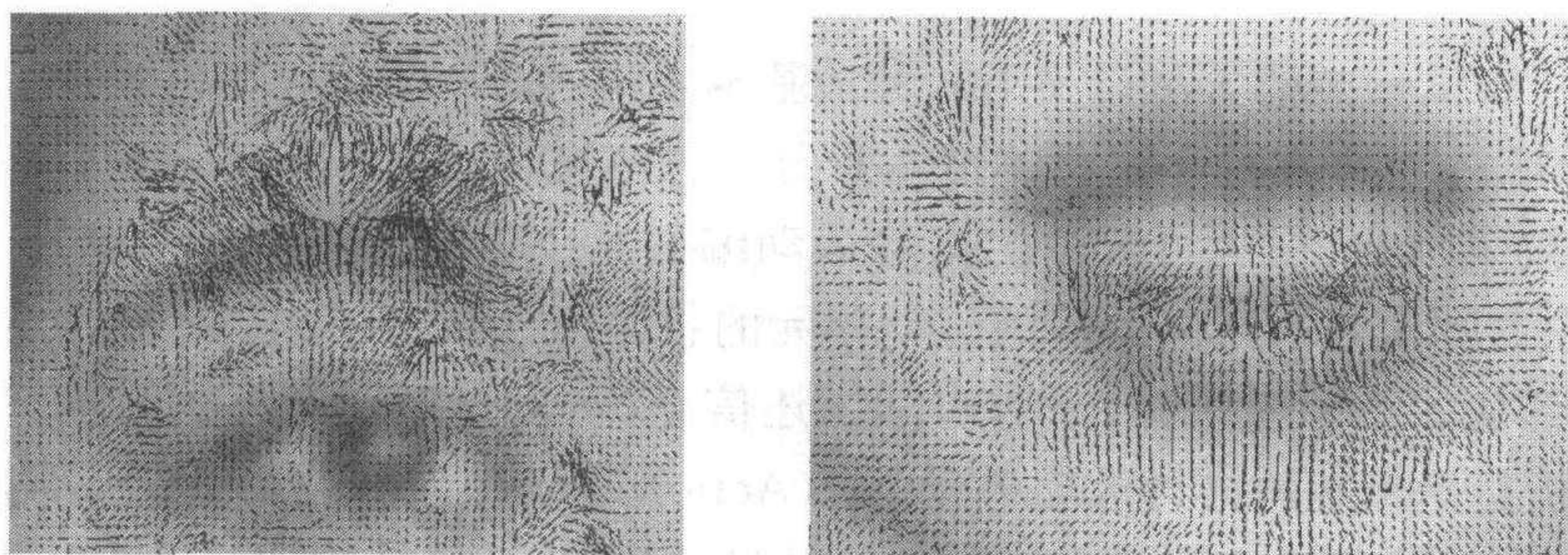


图 2 惊奇表情特征区域的速度场

5 基于 KL 变换^[8]的特征向量的选择

在求特征序列时, 把一组图像序列中的每一帧的不同特征区域的运动向量场向其对应的基底分别做投影, 把 u 分量和 v 分量的系数串行合并作为一个特征流向量. 然后把一组图像序列的特征流向量组成一个特征序列, 如图 3.

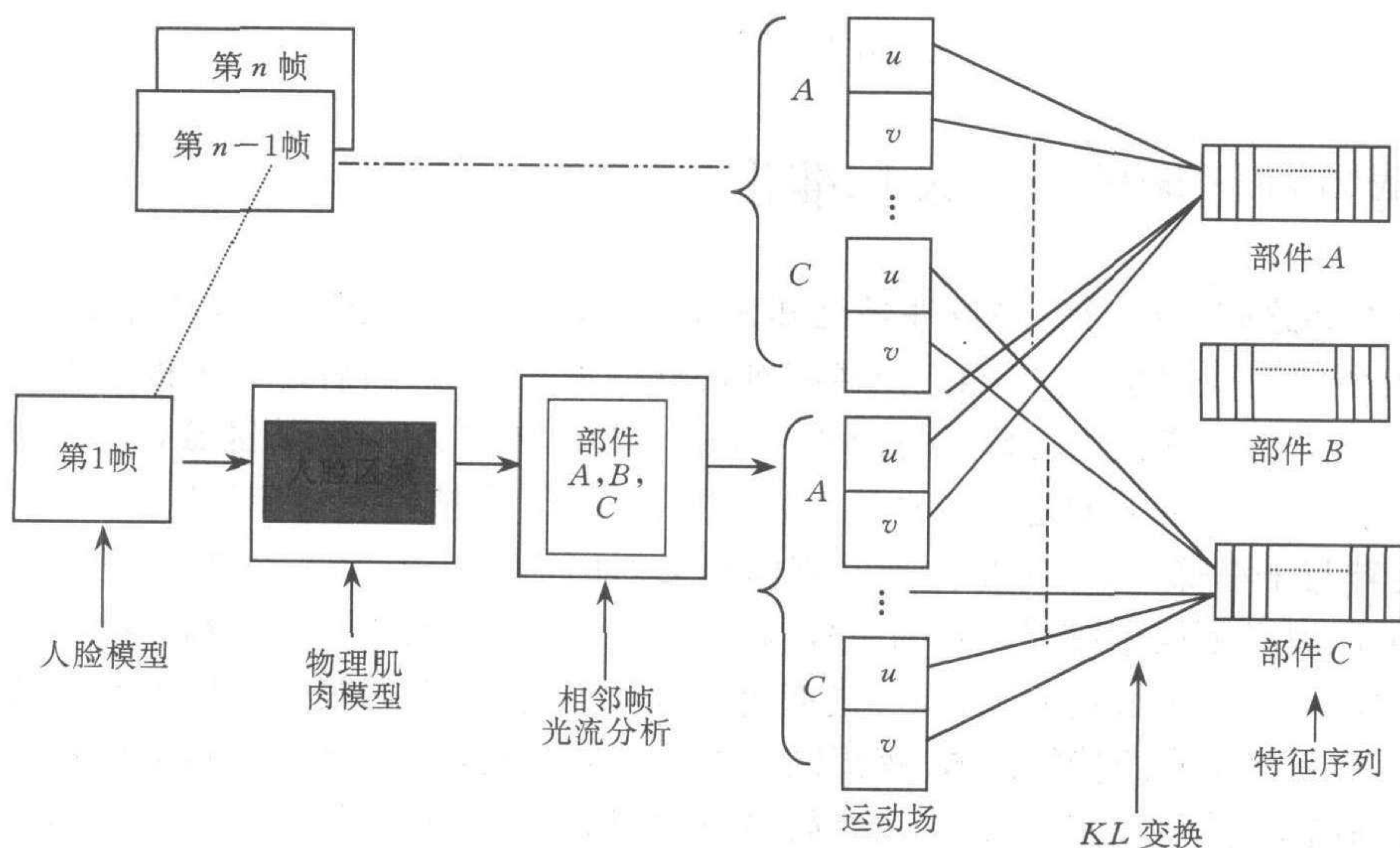


图 3 特征序列的获取

6 基于 HMM 的表情识别

隐马尔柯夫模型^[9]能够解决时变的非平稳信号或过程的模型化问题,用概率或统计范畴的理论成功地解决了怎样辨识具有不同参数的短时平稳的信号段,它既可描述瞬态的随机过程又可描述动态的随机过程的转移特性.由 HMM 所描绘的时间序列的过程是一个状态转移模型.人的面部表情图像序列同人的语音序列特点一样,是有时间顺序的、不可逆过程,因而采用无跨越从左向右的 HMM 模型.在此模型中限定一起始状态,每个状态只能向右侧编号高一位的状态或本状态转移,因此在这一模型的矩阵 A 中只有主对角元素和右副对角元素允许非零.而且 A 比较稀疏,因此大大减少了模型参数估值的计算量.在本实验中采用的是半连续的 HMM,即在每个状态下的码本均值和协方差矩阵是相同的,能够满足计算中的精度和计算量的要求.本实验在 HMM 求解中,原始的光流数据缩小至原来的 $1/20$,避免造成溢出.

7 实验结果

我们用七种最基本的表情做实验,即中性,高兴,惊奇,愤怒,悲伤,厌恶和恐惧.对每种表情采集了 20~30 组图像序列,共 196 组,采集速度为每秒钟 7 帧,大小为 256×256 ,每组的帧数各不相同,大约有 10 多帧.实验中,用人脸面部结构特征的先验信息和灰度积分特征,确定表情特征区域;并对其进行归一和标准化,得到眼部区域大小为 20×30 ,嘴部区域大小为 30×40 ;然后对每组序列的每帧图像进行时间-空间上的平滑,用基于梯度的光流方法提取特征部件区域的运动信息,得到水平方向 u 和垂直方向 v 的运动图像;在对运动特征进行降维时,把所有组所有帧的图像作为训练基底,分别对 u 和 v 方向以及分别对各特征部件进行降维,各部件的特征向量的长度的确定是分别取 u 和 v 基底中较大维数作为长度的一半,各部件的特征向量是 u 和 v 分量上分别投影后串行连接起来的.眼睛部件区域, u 分量和 v 分量按相同的百分比分别可降到 16 和 18,所以取 18 作为 u, v 基底的维数,这样眼睛特征向量长度为 $18 \times 2 = 36$;同理可得到嘴部特征向量的长度为 $20 \times 2 = 40$.

在对 HMM 进行训练时,我们划分眼部七种、加上嘴部七种共 14 种 HMM.把各种表情的眼部与嘴部分别开训练,是因为在实际生活中人的面部表情并不是单一的某种表情,随心情和情绪一样是混合复杂的,表现在面部表情上,不同的特征部件所包含的表情信息可能是不同的.对不同特征区域进行分析,进而分析混合表情的含义.

HMM 在分析时序过程中,速度和识别结果都是非常理想的.与静止的单帧表情图像识别相比,动态时序分析的 HMM 的表情识别方法不论在理论与实验上都取得了显著的效果.特别地,悲伤与厌恶的识别,在静止图像识别中误识率很高,而动态时序分析识别的效果就明显好.

参 考 文 献

- 1 赵力庄. 面部表情的分析与识别[硕士学位论文]. 哈尔滨:哈尔滨工业大学, 1997, 7:14~23
- 2 Mark Rosenblum, Yaser Yacoob, Larry Davis. Human emotion recognition from motion using a radial basis function

- network architecture. In: IEEE Workshop on Motion of Non-Rigid and Articulated Objects, Austin: TX, 1994. 43~49
- 3 Irfan A Essa, Alex P Pentland. Coding, analysis, interpretation, and recognition of facial expressions. In: M. I. T. Media Laboratory Perceptual Computing Section Technical Report No. 325, 1995. 19~20
- 4 James J Lien *et al.* Automated facial expression recognition based on FACS action units. In: Proc. Third International Conference on Automatic Face and Gesture Recognition, 1998. 390~395
- 5 Paul Ekman, Wallace V Friesen. Facial Action Coding System. California: Consulting Psychologists Press Inc., 577 College Avenue, Palo Alto, 94306, 1978
- 6 周昊煜. 脸部图像识别系统的研究与实现[硕士学位论文]. 哈尔滨:哈尔滨工业大学, 1996. 37~40
- 7 高文, 陈熙霖. 计算机视觉. 北京:清华大学出版社, 1999. 84~87
- 8 边肇祺. 模式识别. 北京:清华大学出版社, 1998. 203~206
- 9 Baum L E, Petrie T, Soules G, Weiss N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals Math. Statistics*, 1970, **41**(1):164~191

金 辉 1993年进入哈尔滨工业大学攻读硕士学位, 1996年直读博士学位, 2000年3月博士毕业. 研究领域为模式识别、图像处理.

高 文 教授, 博士生导师. 主要研究领域为多媒体数据压缩、图像处理、计算机视觉、多模式接口、人工智能、虚拟现实等.

《自动化学报》征稿简则

一、《自动化学报》是中国自动化学会主办的高级学术期刊, 每年出版六期.

二、本刊刊载自动化科学与技术领域的中文和英文撰写的高水平学术论文和科学研究成果. 内容包括: 1. 自动控制理论; 2. 系统理论与系统工程; 3. 自动化技术及其在国民经济各领域中的创造性应用; 4. 自动化系统计算机辅助技术; 5. 机器人与自动化; 6. 人工智能与智能控制; 7. 自动控制系统中的新概念、新原理、新方法、新设计; 8. 信息理论与信息处理技术, 模式识别; 9. 自动化学科领域的其它重要问题.

三、本刊发表的文章以论文和短文两种形式为主, 并不定期地发表综述与评论性文章、长论文、书刊评论、问题讨论、读者来信和国内外学术活动信息等.

四、本刊原则上只发表原始性稿件, 但不排除刊登已在国内外学术会议上发表或准备发表的优秀论文的可能性(对于此种情况, 作者必须如实说明).

五、作者投稿时需签署“作者承诺”.

六、来稿格式及要求

1. 来稿要求论点明确、论证充分、语言通顺、文字简练. 一般定稿时论文不超过 6000 字; 短文不超过 3000 字; 其它形式文章视具体内容由编辑部决定. 对重要成果进行系统、完整叙述的长论文字数可稍长. 长论文稿件, 作者在投稿时必须注明.

2. 稿件首页应包括下列内容: 标题; 作者姓名、工作单位、详细通讯地址(包括邮政编码)、E-mail、电话号码; 论文摘要; 关键词; 用英文书写的上述内容.

3. 论文和短文的文章结构请参照本刊最近发表的文章格式. 论文摘要在 200 字以内; 短文 100 字左右. 文中缩写词(中文或英文)须在首次出现时注明全称; 公式、图、表均须分别用阿拉伯数字全文统一编号.

4. 计量单位一律用国际单位, 即 SI 单位制. 名词术语必须规范化、标准化, 前后一致. 外国人名、地名、书刊名称除已通用者外一律用原文.

5. 参考文献按文中出现的先后次序排列, 文献如为期刊时, 按编号, 作者(姓在前如 Wiener L N, Kalman R E, Wang H.). 文章题目, 期刊名(外文可根据国际惯例使用缩写词), 年份, 卷号(期号); 页码顺序编排. 文献如为图书时, 则按编号, 作者(姓在前), 书名, 版次(初版不写), 出版地点: 出版者, 年份, 页码顺序排列. 文中未引用的文献不得列入参考文献栏目.

(下转第 653 页)