

变形模型技术的发展

熊丽华¹, 宋京²

(1. 广东水利电力职业技术学院 计算机系, 广东 广州 510635;

2. 华南理工大学 计算机科学与工程学院, 广东 广州 510640)

摘要: 电影业中有一种有趣的变形动画技术, 它能够通过两幅图像合成出平滑的变形动画特效。近年来, 在模式识别领域出现了一种更为强大的变形技术—变形模型技术。这种技术使用计算机来模拟人类在识别物体时的学习过程, 使用数学表示来建立某对象类的变形模型, 给对象确认、图像合成、图像压缩等计算机视觉领域提供了新的解决方法。

关键词: 线性对象类; 变形模型; 图像表示; 图像分析; 图像合成

中图分类号: TP391.41

文献标识码: A

文章编号: 1673-629X(2006)08-0179-03

Development of Morph Model Technology

XIONG Li-hua¹, SONG Jing²(1. Computer & Infor. Eng. Dept., Guangdong Tech. Coll. of Water Resources
and Electric Eng., Guangzhou 510635, China;

2. Coll. of Computer Sci. & Eng., South China Univ. of Tech., Guangzhou 510640, China)

Abstract: There is a very interesting morph-animation technology in film industry. It can synthesize fluid transformation of one digital image into another. In recent year a more powerful image morphing appeared. The technology simulates the procedure in human learning. It uses mathematical expression to build certain object class, and can provide novel solutions to several vision tasks, including object verification, image synthesis and image compression.

Key words: linear object classes; morph model; image representations; image analysis; image synthesis

0 引言

神经生物学的研究发现, 人类识别物体的能力是通过学习获得的, 学习可以使得大脑视觉皮层中神经细胞间的连接有选择地发生改变, 其结果是形成各种各样的信息分布。图1所示的黑盒模型说明了人类的三种学习过程。第一个黑盒称之为监督学习过程, 首先对给定的已知输入和输出样本进行训练, 然后推导出黑盒中的变换函数 f , 并证明当给出不包含于训练集中的同类输入图像时, 能由 f 推导出正确的输出。第二个黑盒表达的是图像分析过程, 输入是一组图像, 输出是一组与输入图像相关的参数, 这些参数可以被看作是输入图像的特征, 如物体的相对坐标、三维结构、纹理属性等。第三个黑盒是第二个黑盒逆向过程或延伸, 即通过参数信息来恢复已知图像或者是合成新的未知图像。文中介绍的变形模型技术就是通过模拟人类的学习过程, 来实现用机器理解图像、认知图像, 反过来加深人类对自身识别图像认知过程。

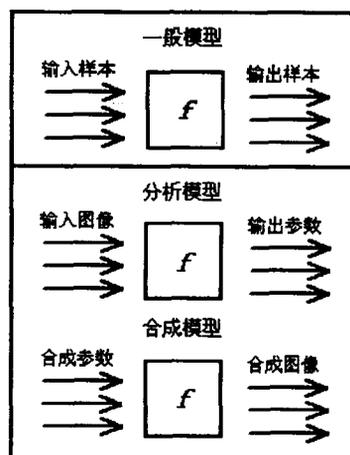


图1 黑盒模型

1 变形模型的基本概念及实现步骤

图像变形的概念源于电影业中变形动画技术^[1]的应用。首先选定两幅图像, 然后合成若干张中间平滑过渡图像, 最后制作出平滑的变形动画特效。文中介绍的变形模型技术更为强大, 融合了多种学科的新技术。因为变形合成参数是多维的, 所以使用这种技术所建立的对象类模型, 称之为多维变形模型。文中只以人脸类为研究对象来

收稿日期: 2006-03-05

基金项目: 广州市天河区科技计划项目(2002X GP11)

作者简介: 熊丽华(1978-), 女, 湖北咸宁人, 助教, 硕士, 研究方向为空间数据库、图像处理与模式识别。

加以讨论。

多维变形模型的思想源于线性类^[2]的概念,线性类的思想是所有变形模型算法的基础。它认为通过对少量样本图像的学习,能够找到一个线性关系表示出某一类对象,进而可以用线性变换估计出变形图像。公式(1)是线性变换数学表达式, $T_i(x,y)$ 是纹理向量, $S_i(x,y)$ 是形状向量, (x,y) 是像素点。

$$S = \sum_{i=1}^q a_i S_i, T = \sum_{i=1}^q \beta_i T_i \quad (1)$$

Michael Jones 和 Tomaso Poggio 最早提出了多维变形模型的思想^[3],详细阐述了多维变形模型技术的实现过程,包括 3 个主要步骤:图像表示(匹配)、图像分析、图像合成^[4],对应了黑盒模型中的分析模型和合成模型。

图像匹配是一项重要的预处理工作,图 2 是笔者使用光流算法得到的合成图,算法的目的是找到所有对应点的匹配关系。现在假设要寻找与人脸 A 鼻尖上的 a 点相对应的人脸 B 鼻尖上的 b 点,如果直接将两幅样本人脸 A, B 的各点像素值取 1/2 后相加,就会得到等号右边上图的结果,图像显得有些模糊,在某些特征比较明显的部位如眼睛和嘴巴处,可以清楚地看到重影效果。如果将样本人脸 A, B 在进行匹配后再合成,就能够得到如等号右边下图所示的清晰图像。

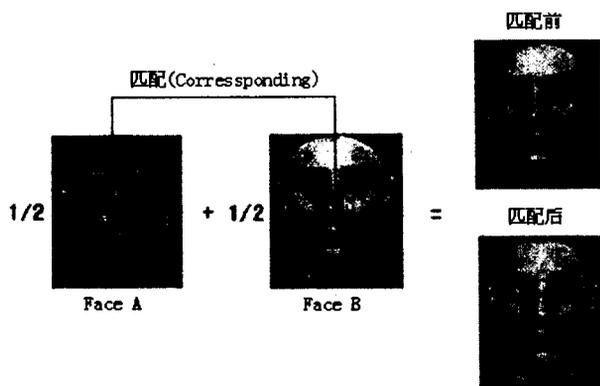


图 2 比较匹配和不匹配时的合成图像

图像分析是变形模型技术的核心,其目的是要找到样本类与给定的一张属于样本类的未知图像之间存在的数学关系,采用图像分析技术计算出公式(1)中的线性样本类系数。图像合成则是把这些系数回代到公式(1)中,从而得到一幅合成图像(如图 3 所示)。为了使得变形模型图像与新样本图像尽可能相似,需要采用一种评判策略,策略之一就是计算最小化误差函数。公式(2)是最小化误差函数,它是变形模型理论中的一个重要公式,使得 E 最小的一组参数值,即为使得变形模型图像与新样本图像最为相似的最佳合成参数。

$$E_I(\alpha, \beta, \dots) = \sum_{x,y} \| I_{\text{input}}(x,y) - I_{\text{model}}(x,y) \|^2 \quad (2)$$

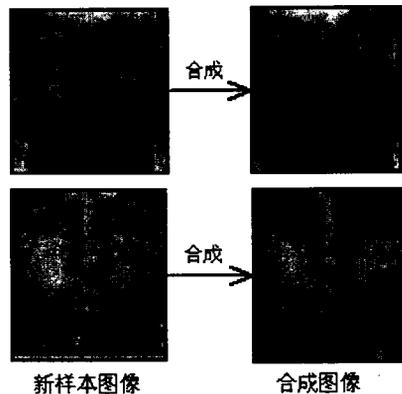


图 3 MIT 人脸数据库图像的合成效果图

2 变形模型主要技术的发展

为了能让变形模型理论有效地应用于图像分析的各种领域,必须对其 3 个主要步骤进行算法优化。下面这些算法分别在变形模型的表示、分析与合成的准确性、健壮性以及计算时间上进行了优化:

① 匹配阶段:光学流匹配算法、由粗到精的匹配法、步步为营的匹配法、PCA 线性类法。

② 分析合成阶段:由粗到精的合成法、分层的合成法、随机梯度下降的最优化算法。

2.1 匹配算法

建立变形模型的核心步骤是确立参照人脸图像与其它样本图像间密集的对点匹配关系,光学流算法能够解决这一问题。普通光学流算法能够在假定图像序列的亮度保持恒定的前提下,找到与 $I_0(x,y)$ 点相匹配的 $I_j(x + \Delta x, y + \Delta y)$ 点,但是不连续图像(人脸图像)间有较大的差异,亮度恒定的假设不再成立,所以必须采用对传统的光学流方法加以改进。Bergen 和 Hingorani 提出了一种由粗到精的光学流计算方法^[5]。它首先使用高斯金字塔生成若干级的高斯金字塔图像,然后在最高级即最小的图像上计算光学流,计算出的结果将作为下一级计算光学流的参考值,逐级递推,直到最后一级图像即原图上的光学流被计算出来。

因为图像匹配是预处理工作,只需要在构造模型阶段计算一次,所以不必考虑算法的速度问题。Thomas Vetter 提出了一种步步为营的匹配算法^[6],在光学流算法失效的情况下,它可以得到很好的匹配效果。这种方法以从少量样本中合成出来的粗糙变形模型为基础,通过向粗糙变形模型不断添加新样本,逐渐提高模型的合成能力。另一种等效的步步为营方式是通过增加选取的 PCA 主成份数来不断强化变形模型。基于样本集的平均样本值与各样本的距离最小的原理,选择平均图像作为参照图像,能够有效地提高匹配的准确性。

2.2 分析合成算法

合成新样本图像的核心是计算最小化误差函数,计算方法用到了数值优化算法。解决数值优化问题有许多经

典的算法,为了有效地提高收敛速度和避免陷入局部最小,选择的是随机梯度下降算法^[7]。随机的思想体现在随机地从图像中选取若干像素点,只计算所选择点的梯度。为了提高合成系统的健壮性,由粗到精的思想也被引入到合成算法中。

为了进一步提高合成能力,Michael Jones 和 Tomaso Poggio 提出了层级变形模型^[7]。它采用把样本图像切分成较简单的独立组件的思想,对于人脸类图像,这些独立的组件包括眼睛、鼻子和嘴巴。眼睛、鼻子和嘴巴是低一层级的变形模型组件,它们中心点的线性组合构成了高一层级的变形模型,组合这两层变形模型就可以合成出更优的新样本。因为组件远没有整张图像复杂,所以只需要较少的样本就可以表达出组件中的信息,而且合成质量与标准变形模型相当。此外,层级模型对部分遮挡的图像有着更好的处理能力,当一个组件由于被遮挡而不可见时,其它组件仍能够毫不受影响进行合成,虽然标准变形模型也有这种能力,但是它最后合成的图像的质量会随着遮挡范围的增大而不断下降。

为了优化迭代算法,Xun Xu 和 Changshui Zhang 提出了一种灵活的变形模型技术^[8]。他们认为随机梯度下降算法没有充分考虑到待合成图像中隐含提供的信息,在合成新样本时效率不够高。灵活变形模型是利用新样本中隐含的启发信息,提出的一种特殊的迭代算法。在原始变形模型的合成数学表达式中,所有参数都在同一个误差函数里,这使得这些参数很难有效地被更新。考虑到形状和纹理彼此独立这个假设,灵活变形模型先将式(1)分解成两个独立的误差函数,然后通过一些数学变换转化成两次求解最小二乘法的问题。这种算法对初值的选取不敏感,迭代次数更少,迭代速度更快。

2.3 三维变形模型方法

Volker Blanz 和 Thomas Vetter 最早提出了三维变形模型的理论^[9],其思想直接源于二维变形模型技术。这种三维变形技术最具特色之处是它的输入样本不是一幅未知的三维人脸,而是一张未知的普通二维人脸图像。

因为二维人脸图像提供的信息量很少,所以当人脸的类内属性(如表情、年龄)或者是类外属性(光照)发生较大变化时,要想合成出理想的新样本几乎是不可能的,这也限制了二维变形模型技术的应用。因为三维表示能够准确并且独立地表示出任何光照、姿态和表情等属性,所以可以使用三维样本来替代二维样本。三维变形模型同样使用了线性类思想,未知新样本可以由已知的一定数量的三维样本通过线性组合表示出来。光学流算法在进行改进后,也用在了三维图像的匹配过程中, $I(h, \Phi) = (r(h, \Phi), R(h, \Phi), G(h, \Phi), B(h, \Phi))$ 表示了三维人脸的点, (h, Φ, r) 是三维人脸柱面坐标, R, G, B 是三维顶点的颜色值。此外,为了提高匹配的效果,在光学流算法失败时,由粗到精的方法也被采用。

三维变形模型使用 PCA 算法构建线性类,如公式

(3),有两个优点,其一,可以起到降维的作用,以减少合成时的计算量;其二,PCA 也是一种空间变换的统计分析算法,它提供了人脸空间的概率密度估计,能够用来控制合成人脸的真实度。

$$S = \bar{S} + \sum_{i=1}^{N_s} \alpha_i \cdot S^i, T = \bar{T} + \sum_{i=1}^{N_t} \beta_i \cdot T^i \quad (3)$$

在合成阶段,三维图像的变换比二维图像要复杂得多,主要包括三维图像的坐标变换和光照颜色变换。坐标变换可分为三维仿射变换和从三维空间到二维空间的投影变换;光照颜色变换可分为光照变换和颜色变换。为了减少建模样本数,二维层级变形模型的思想也被引申到三维。图像合成在三维变形模型中也叫做“拟合”。目前的三维拟合算法还没有做到完全自动化,必须手工在变形模型和新样本上指定若干特征点,用做迭代时的初始信息。为了有效地控制拟合过程向理想的方向进行下去,使用了一种叫做 MAP 的最大后验概率方法^[10]。

3 变形模型的应用及发展前景

人类识别能力的形成需要一个学习过程,学习是为了发现物体的特征信息。多维变形模型技术正是从这一发现中提出的建模技术,它认为任何一类物体都能够通过对少量的样本的学习之后,用线性关系表示出来。虽然二维变形模型在实验中有着不错的合成效果,但是前提是新的未知图像必须与样本图像具有基本相同的属性,如相同视角或相同光照。只有通过增加具有各种属性的新样本,才有可能有效地合成相应属性的未知图像。因为三维样本克服了二维样本的信息量少、不能真实地表示出物体自身属性和外部环境因素的缺点,所以在二维模型技术的研究成果的基础上,三维变形技术便具有了更为强大的合成能力。

只需要提供一张同类别的二维图像,就可以构建出这个物体的三维模型;反过来还能够利用这一模型来合成各种各样的新图像,这些新图像还能够进一步被用在其它的图像分析领域,用来解决各种由于图像属性变化造成的疑难问题。在图像识别领域中,光照和姿态的变化一直是影响识别率的重大因素,可以借助三维变形模型技术,把所有图像的光照和姿态变换到标准状态后再进行识别。

三维变形模型技术有几个方面有待进一步发展以加强整个系统的性能。第一,反射和光照问题。当前的拟合方法只能使用光照锥模型恢复出单个直射光源,假如使用的是更加接近人脸皮肤属性的反射以及更接近照片中人脸所处环境的光照,这将有效地提高三维图像的重建效果。第二,没有实现完全自动化的合成过程,仍需要手工选择特征点。这一点可以借助于特征自动定位技术加以改进。第三,目前用到的迭代算法除随机梯度下降之外,还有随机牛顿下降算法、ICIA 算法(反向合成图像对齐算法),但是它们的拟合速度都不够理想,不能满足实际应用

(下转第 185 页)

数据库设计在涉及到后台数据库的动态网页的开发中是至关重要的一个方面,其设计的好坏直接影响系统的执行效率以及日后的维护,还应考虑到良好的扩展性要求。通过需求分析,笔者设计在 SQL Server2000 中建立一个数据库,其中包含用户信息表、管理员信息表、试题信息表(其数量根据科目而定)、题目类型表、试题难度表和用户反馈信息表。

现在详细说明一下最主要的试题表的所有字段:

①题号(主键)。使用数据库自动生成的 ID;

②题型。这是来自题型表的一个外键,如:题型代码“2”对应于页面上显示的具体信息——“选择题”,这样可以加快查找试题的速度;

③难度。与题型类似,在此填写的值是难度表中的难度代码,如:难度代码“2”对应于页面上显示的具体信息——“较容易”;

④分值。用户可以根据需要,通过修改后台数据库中的分值以实现不同的分值;

⑤主题。用于查询试题时的模糊匹配查询,以找到题干与主题匹配的试题;

⑥题干。即试题的题目;

⑦答案。在进行自动组卷后,系统应根据所选择的题目给出相应的试题答案。

数据库中最主要的试题表的具体内容如表 1 所示。

4 结束语

远程教育 COD 系统以流媒体课件为基础,针对学生的起点不同,本系统还具有自适应性,即系统可根据学生的个别能力特征,动态呈现与学习者当前学习能力最相关

的基于超媒体的教学内容。同时,虚拟实验室和作业管理功能可以巩固学生从课件中学习到的知识。在线考试系统可以用来对学生进行考核,以此为据决定是否给予学生该门课程的学分。互动模块则为师生提供一个相互交流的平台,使得学生在学习过程中遇到的问题能够及时地反馈给老师,老师也可以通过互动平台在线解决问题,从而提高网络教学的效率。

表 1 试题表

字段名	字段描述	字段类型	字段长度
TNo	题号	int	6
Ttype	题型	int	2
Diffi	难度	int	2
Point	分值	int	2
Topic	主题	varchar	200
Body	题干	varchar	1000
Answer	答案	varchar	1000

参考文献:

- [1] 周霖,武祥村,蒋东兴,等.课件点播系统研究与实现[EB/OL]. <http://218.65.59.6/kejiandb/kejianjs/yanjiusx.htm>,2005-03-17.
- [2] 何丰如.基于 Web 的智能化网络教学系统的结构分析与设计[J].广东广播电视大学学报,2003(3):24-30.
- [3] 蒋东兴,张继才,罗念龙.高校网络教学现状与对策[J].计算机教育,2004(9):44-46.
- [4] 孙三才,许薰尹.精通 C# 与 ASP.net 程序设计[M].北京:中国铁道出版社,2003.
- [5] 冯建华,周立柱.数据库系统设计与原理[M].北京:清华大学出版社,2004.

(上接第 181 页)

的要求,所以应该充分利用图像中包含的先验信息,设计出迭代速度更快的优化算法。

参考文献:

- [1] Wolberg G. Image Morphing - A Survey[J]. Visual Computer, 1998(14):360-372.
- [2] Vetter T, Poggio T. Linear Object Classes and Image Synthesis from a Single Example Image[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1997,19(7):733-742.
- [3] Jones M, Poggio T. Model-based matching of line drawings by linear combinations of prototypes[C]. USA: MIT, 1995. 531-536.
- [4] Jones M J, Poggio T. Multidimensional morphable models [M]. ICCV: IEEE Computer Society Press, 1998. 683-688.
- [5] Bergen J R, Hingorani R. Hierarchical Motion-based Frame Rate Conversion[R]. Princeton, USA: David Sarnoff Re-

search Center, 1990.

- [6] Vetter T, Jones M J, Poggio T. A bootstrapping algorithm for learning linear models of object classes[M]. CVPR, Washington DC, USA: IEEE Computer Society, 1997. 40-46.
- [7] Jones M, Poggio T. Hierarchical Morphable Models[A]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C]. Santa Barbara, CA, USA: [s. n.], 1998. 820-826.
- [8] Xu Xun, Zhang Changshui, Huang T S. Active Morphable Model: An Efficient Method for Face Analysis[A]. The 6th International conference on Automatic Face and Gesture Recognition[C]. Seoul, Korea: [s. n.], 2004. 837-842.
- [9] Blanz V, Vetter T. A Morphable Model for the Synthesis of 3D Faces[A]. In: Rockwood A. Proc. of the SIGGRAPH'99[C]. New York: ACM Press, 1999. 187-194.
- [10] Blanz V, Vetter T. Face Recognition Based on Fitting a 3D Morphable Model[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003,25(9):1063-1074.