

多源类比生成在影像序列生成中的应用^{*}

劳志强 徐冬溶 潘云鹤

(浙江大学 CAD&CG 国家重点实验室, 人工智能研究所 杭州310027)

摘要 本文介绍了多源类比生成的理论,并详细描述了我们的基于图形学和多源类比生成理论所进行的一个影像序列生成的过程。

关键词 影像生成技术, 类比生成, 多源类比

一、前言

计算机辅助影像生成技术在近十年来取得了飞速的发展。许多成功的影像制作系统,如,美国的 WAVEFRONT, 加拿大的 SOFTIMAGE 和 ALIAS, 法国的 TDI, 已被成功地应用于商业广告、电影电视特技、游艺场所以及教育、军事、科研甚至于法院案件的审理。近几年,浙江大学、北方工大、中科院软件所等单位相继从国外引进了商用影像制作软件,并在电视片头、影像广告设计等方面取得了很好的成效。浙江大学用 ALIAS 和 SOFTIMAGE 进行了古兵马俑的影像设计;中科院软件研究所用 TDI 的 EXPLORE 设计了第十一届亚运会的片头;计算机辅助影像生成技术将影响着人们未来的生活方式。

尽管现在影像生成技术方面的研究已经取得了令人可喜的成果,基于这些成果的商用影像系统在物体建模、运动控制及图形绘制方面有很强的功能,但这些系统存在着许多局限性,比如说,在拍摄电影时,导演在执导剧本和演员在表演过程中能够控制整个演出过程,而目前的影像系统还不能适合动态的设计环境。这里,我们试图把知识处理技术应用到模型影像中,使影像技术在处理过程中接近于人的行为。

N. M. Thalmann 和 D. Thalmann 把影像生成技术分成五级,最高级为智能影像^[1]。传统影像技术对硬件设备要求很高,因为任何一帧影像都是基于模型计算出来,影像序列生成过程中的计算量非常之大。另外,在计算机广告制作以及电影特技中,创意是独立于影像制作之外的,是影像设计者结合以往的知识所进行的形象思维的体现,如何才能将人的这部分思维全部或部分地融入影像技术中去呢?我们想用人工智能领域中的类比生成技术在这方面

做些尝试。

类比是人类学习和思维的一种重要方式,随着近年来对机器学习问题研究的深入,对类比的研究也愈来愈被人们所重视。类比的研究始于心理学^[2-4]、哲学^[5]等领域,而近年来的类比研究热潮却在心理学和人工智能领域,Corbonell, Gentner, Gick, Winston 等人的一批重要文献^[7-10]为类比推理研究定下了基调,他们提出了一批类比机模型如: SME, ACME, IME 等,将类比的研究分为 Syntactic 和 Pragmatic 两类^[11],我国学者在这方面也进行了卓有成效的研究工作^[12-14],并已成功地将部分应用于自动程序设计等方面。

类比和演绎、归纳相比是一种约束更为宽松的推理模式,只要求推理对象之间存在某种程度的相似,结果是似然性的,亦非永真。类比生成是一种新的推理方法,与归纳、演绎一样在人类解决问题中大量被采用。在创造性的设计领域,经验丰富的设计师在面对一个新问题时,总是联想到一大批有用的素材,通过加工整理,综合多个类比源生成最终的设计结果。

多源类比生成与普通类比推理不一样。通常的类比基于假设:事物 A 和事物 B 在某些方面的相似性表明它们在另外一些方面也相似,而多源类比生成不存在这样的假设,它是一种生成性的过程,在这个模型下的所有类比源,都被认为是“广义相似”的,因而可以被用来生成所有源都感兴趣的结果。同时,传统的类比基于符号,因而是二值性的,而多源类比生成是面向形象的,可以在连续函数上实现,而且,多源类比突破了传统类比仅能进行两个源类比的限制。基于多源类比生成理论,浙江大学人工智能研究所完成了国家八五攻关项目“彩色拼版制版 CAD/CAM 系统 ADS2.0”,主要面向广告设计师。该系统已通过电子工业部主持的专家鉴定,获得了较好的

^{*} 本文得到国家自然科学基金资助

评价(国际先进)。目前我们在此基础上进行将多源类比生成理论用于影像序列生成的研究。本文就我们前段时间所进行的研究工作做一个小结。

二、多源类比生成模型

设 $B_1, B_2, \dots, B_N$ 为 N 个类比源, T 为靶, 各有 M 个部件组成, 则:

$$B_i = \{b_{ij} | j=1, 2, \dots, M\}, i=1, 2, \dots, M_1$$

$$T = \{t_j | j=1, 2, \dots, M\}$$

传统类比推理为: $M=2, N=1$, 则有: $t_2 = f(B_{11})$ 。现去其约束, 则有:

$$t_1 = f_1(B_{11}), t_2 = f_2(B_{12}) \quad (2.1)$$

上式是基于单源的, 现扩充其源数量, 则有:

$$T = F(B_1, B_2, B_3, \dots, B_N) \quad (2.2)$$

$$\text{令 } A = \begin{bmatrix} a_{11} & \dots & a_{1M} \\ a_{21} & \dots & a_{2M} \\ \dots & \dots & \dots \\ a_{N1} & \dots & a_{NM} \end{bmatrix}_{N \times M} = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_N \end{bmatrix},$$

$$B = \begin{bmatrix} b_{11} & \dots & b_{1M} \\ b_{21} & \dots & b_{2M} \\ \dots & \dots & \dots \\ b_{N1} & \dots & b_{NM} \end{bmatrix}_{N \times M} = (B_1, B_2, \dots, B_N)$$

$$\left\{ \begin{array}{l} T = \{t_1, t_2, \dots, t_M\} = \sum A_i B_i, \\ A_i = \begin{bmatrix} a_{i1} & 0 & \dots & 0 \\ 0 & a_{i2} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{iM} \end{bmatrix}_{M \times M} \end{array} \right. \quad (2.3-1)$$

$$\sum A_i = I \quad (2.3-2)$$

$$\left\{ \begin{array}{l} a_{ij} \in \{0_F, 1_F\} \subset F \\ a_{ij} \in [0_F, 1_F] \subset F \end{array} \right. \quad (2.3-3)$$

或:

$$\left\{ \begin{array}{l} T = BA = (B_1, \dots, B_N) \begin{bmatrix} a_{11} & \dots & a_{1M} \\ a_{21} & \dots & a_{2M} \\ \dots & \dots & \dots \\ a_{N1} & \dots & a_{NM} \end{bmatrix} \\ \left\{ \begin{array}{l} a_{ij} \in \{0_F, 1_F\} \subset F \\ a_{ij} \in [0_F, 1_F] \subset F \end{array} \right. \end{array} \right. \quad (2.4)$$

式(2.3-1)中的 A_i 为相似度量, 定义 T 与 B_i 的相似程度, (2.3-2)保证 T 与 B_i 的规模一致, (2.3-3)式保证 T 与 B_i 或多或少相似(狭义相似)。

设 $B_i, T \in V, \oplus$ 对 V 封闭, 则 $(V, \oplus)$ 构成代数结构, 是可交换群, $\sum$ 是 $\oplus$ 的缩写:

$$\textcircled{1} t_1 \oplus (t_2 \oplus t_3) = (t_1 \oplus t_2) \oplus t_3, \text{结合律}$$

$$\textcircled{2} \exists 0 \in V, \forall x \in V, \text{则 } 0 \oplus x = x \oplus 0 = x$$

$$\textcircled{3} \forall x \in V, \exists y \in V, \text{使 } x \oplus y = y \oplus x = 0$$

$$\textcircled{4} t_1 \oplus t_2 = t_2 \oplus t_1, \text{交换律}$$

V 是域 F 上的向量空间, 即: $\textcircled{1} (V, \oplus)$ 为可交换群. $\textcircled{2} \forall a, b \in F, (a \oplus b)x = ax \oplus bx; a(x \oplus y) = ax \oplus by; (ab)x = a(bx), 1_F x = x$.

现在考察二源 B_1, B_2 单部件情况, 以推广(2.3)式($M=1, N=2$), 则有:

$$T = \{t_1\} = A_1 B_1 \oplus A_2 B_2 = a_{11} b_{11} \oplus a_{21} b_{21}$$

$$a_{11} \oplus a_{21} = 1_F$$

$$a_{ij} \in [0_F, 1_F] \quad (2.5)$$

由此, T 为由 B_1 和 B_2 类比生成, 现令上式中 B_1 为未知, T 和 B_2 为类比源, 则有:

$$a_{11}^{-1} a_{11} B_1 = a_{11}^{-1} T - a_{21} a_{11}^{-1} B_2$$

$$a_{11} \oplus a_{21} = 1_F. \text{令 } k_1 = a_{11}^{-1}, k_2 = -a_{21} a_{11}^{-1}$$

$$a_{11}, a_{21} \in [0_F, 1_F]$$

则, $B_1 = k_1 T + k_2 B_2, (a_{11}^{-1} a_{11} = 1_F, 1_F x = x)$

$$k_1 \oplus k_2 = a_{11}^{-1} (1_F \oplus (-a_{21})) = a_{11}^{-1} (a_{11} \oplus (a_{21} \oplus (-a_{21}))) = a_{11}^{-1} a_{11} = 1_F$$

$$k_1, k_2 \in F (k_2 < 0_F)$$

此时上式第一式表明 B_1 为从 T 和 B_2 类比而得, 相似度各为 k_1, k_2 , 因 k_2 为负, 因此 B_1 是从 T 中除去 B_2 的特性所得, 第二式仍用作保证 B_1 的规模与 T 和 B_2 一致, 第三式使类比从局部区域扩为整个 F , 因此得到二源一部件的一般形式为:

$$T = a_{11} B_1 + a_{21} B_2 | a_{11} \oplus a_{21} = 1_F | a_{11}, a_{21} \in F \quad (2.6)$$

同理可证, 当推广至二源 M 部件有:

$$T = A_1 B_1 + A_2 B_2 | A_i \oplus A_j = I_M | a_{ij} \in F | i=1, 2; j=1, 2, \dots, M \quad (2.7)$$

进一步推广至多源多部件有:

$$T = A_1 B_1 \oplus A_2 B_2 \oplus \dots \oplus A_N B_N = \sum_{i=1}^N A_i B_i$$

$$\sum_i A_i = I_M$$

$$a_{ij} \in F \quad (2.8)$$

其中:

$$A_i = \begin{bmatrix} a_{i1} & 0 & \dots & \dots & 0 \\ 0 & a_{i2} & 0 & \dots & \dots \\ \dots & 0 & \dots & 0 & \dots \\ \dots & \dots & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & a_{iM} \end{bmatrix}$$

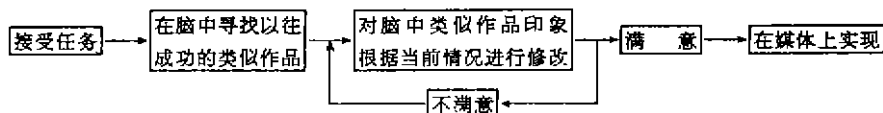
式(2.8)中 a_{ij} 为 T 与第 i 个源第 j 个部件的相似度系数, A_i 为 T 与第 i 个源的相似矩阵, 第一式 A_i 定义 T 与各 B_i 之相似程度, 因而 T 由 B_i 类比生成, 第二式保证 T 与各 B_i 在规模上保持一致, 第三式说明 T 与各 B_i 或相似(为正)或不相似(为负), 即与各 B_i 是广义相似的(相似度可正可负)。

三、多源类比影像序列生成的图形学基础

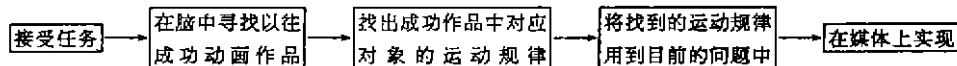
计算机生成的影像实际上是一组图象的序列,

其中当前帧画面是对前一帧的部分修改。当我们拿到一张照片而被告知要做一个照片上对象的影像时,马上会在脑中找寻以往做过的(或见过的)类似的成功影像作品,然后依照所给照片的实际情况对

脑中以往类似的成功影像作品做必要的修改就得到了现在所做的影像作品,再动手将脑中的作品在纸上或其它媒体上展示出来,这个形象思维的过程如下图所示:



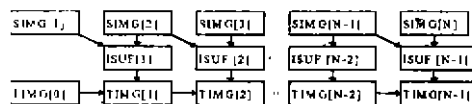
换言之,可如下图所示:



我们知道,在平面上作影像序列生成常常将每一帧图象的前景和背景分开,要动的对象作为前景处理,不动的部分作为背景处理。采用前景、背景分别对待的处理方法,即只关心影像序列生成过程中帧与帧之间的变化,可以大大地提高影像序列生成的速度。这里所用的影像序列生成方法的创新之处

就在于将影像对象与艺术家的感觉融合在一起,即由艺术家在影像的关键帧图象上勾勒出影像对象的特征点、线、曲线,然后再在这些定义好的特征之间进行平滑的插值最后生成符合艺术家想象的影像结果来,这个影像序列生成过程的详细框图如图1所示。(设源图象序列长度为 SOUR_SEQ_LENGTH)

根据上述框图我们得到源与靶之间的对应关系如下所示:



这里 $SIMG[1], \dots, SIMG[N], TIMG[0]$ 为 $N+1$ 个类比源, $IIMG[1], \dots, IIMG[N-1]$ 为由 $SIMG[i] (i=1, 2, \dots, N)$ 产生的插值曲面, $TIMG[i] (i=1, 2, \dots, N-1)$ 为靶, 各有 M 个部件, 这里我们将图象空间中两点间的位移差对 $(\Delta x, \Delta y)$ 定义相似度的度量, 考虑到这个类比生成过程为二源 M 部件, 有:

$$TIMG[i] = A_1 TIMG[i-1] + A_2 ISUF[i]$$

从上式可以看出解决问题的关键在于求得相似矩阵 A_1 和 A_2 。由上面的图形学知识可知, 插值曲面就是相似矩阵的具体表现形式, 代入上式后我们就可以求得 $TIMG[i]$, 同理可求得余下的靶图象及中间图象。

结束语 自1981年我国著名学者钱学森提出思维科学的体系(抽象思维学, 形象思维学, 灵感思维学)以来, 学术界对形象思维的研究进入了一个新的高潮。对于设计领域(如, 建筑, 广告等)中, 如何将人的形象思维信息融入计算机辅助设计中是一个亟待解决的问题。本文在这方面作了一个尝试, 要进一步研究使之实用化, 须与知识库技术结合起来, 建立相应的知识库, CASE 库和素材库。目前我们正在进行更进一步的研究。(参考文献共17篇略)

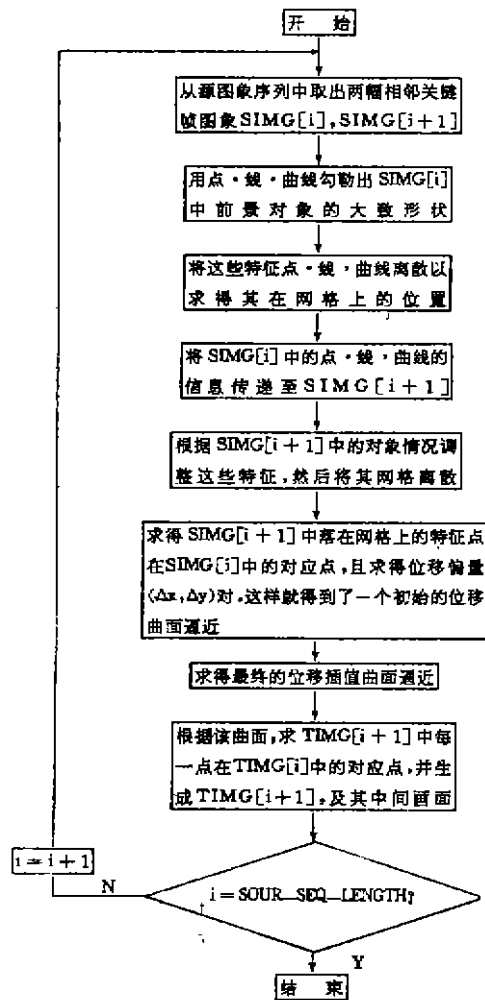


图1