

基于学习的烙画图像特征重构

王 东^{1,2}, 周世生¹

(1. 西安理工大学印刷包装工程学院, 西安 710048; 2. 浙江工业大学之江学院, 杭州 310024)

摘 要: 提出一种基于学习的烙画图像计算机仿真算法。在 Hertzmann 图像类比算法的基础上, 使用融合轮选择算子的粒子群优化算法对处理过程进行加速, 整个粒子群的当前最优位置采用轮选择的方法, 能有效降低粒子群优化算法过早收敛于局部最优的机率, 获得较传统近似最近邻域搜索更快的处理速度。为弥补 Hertzmann 算法中结果图像视觉效果上的纹理缺陷, 在图像融合过程中, 先将 2 幅输入图像转换到 $l\alpha\beta$ 颜色空间, 再进行点对点的加计算, 进而实现图像视觉效果的增强。烙画图像的类比实验结果表明, 该算法所获得的烙画仿真结果图像接近真实烙画图像效果。

关键词: 烙画; 学习; 轮选择; 粒子群优化

Learning-based Feature Reconstruction for Pyrography Painting

WANG Dong^{1,2}, ZHOU Shi-sheng¹

(1. Faculty of Printing and Packaging Engineering, Xi'an University of Technology, Xi'an 710048, China;

2. Zhijiang College, Zhejiang University of Technology, Hangzhou 310024, China)

【Abstract】 This paper presents a novel learning-based approach for pyrography painting style simulation. Based on Hertzmann's image analogies algorithm, it proposes a novel Particle Swarm Optimization(PSO) combined with the roulette selection operator to speed up the image processing. This algorithm selects the best position found by the swarm so far with roulette wheel selection method and the probability of premature converge to local minima is decreased. Compared with the Approximate Nearest Neighbor(ANN) search, it can obtain faster processing speed. In order to compensate for visual defects on the output texture in Hertzmann's algorithm, it converts RGB signals to perception-based color space $l\alpha\beta$ before the additional computing. Experimental results demonstrate the algorithm is efficient for pyrography painting simulation.

【Key words】 pyrography; learning; wheel selection; Particle Swarm Optimization(PSO)

1 概述

近年来, 艺术风格图像的计算机模拟作为计算机图形学一个新的分支为广大图像工作者所关注。目前, 人们已经能够对油画、水彩画、铅笔画、钢笔画、版画甚至中国画实现近似的模拟^[1]。这些算法大致可以归为基于笔触模型和基于艺术风格学习两大类。基于笔触模型的方法需要大量的交互准备工作, 使其应用受到一定限制。文献[2]提出图像类比的方法, 在影像重建、纹理合成与传输、铅笔及油画风格渲染等方面获得了很好的效果, 可拓展性很强。该方法中使用了近似最近邻域(Approximate Nearest Neighbor, ANN)搜索方案。由于 ANN 算法自身性能的原因, 其全局最优解存在一定的局限性, 与人的视觉感受存在一定差距^[3]。一致性邻域搜索则能够保持图像中像素的连续性, 比 ANN 算法获得更好的视觉效果。因此, 在 Hertzmann 算法中采用一致性邻域搜索来弥补 ANN 算法的不足。尽管如此, 由于 ANN 算法采用的是二叉树的数据处理方式, 无法排除大量冗余的像素点, 致使查找的工作量变得极为庞大, 处理时间增长。

本文利用融合轮选择算子的粒子群优化(Particle Swarm Optimization, PSO)算法提高其加速性能。该算法能够从速度与收敛 2 个方面兼顾整个算法的性能, 能有效防止算法陷入局部最优, 并提高粒子群整体的平均适应度。针对 Hertzmann 算法在学习过程中背景纹理处理能力的不足, 本文根据文献[4]提出的 $l\alpha\beta$ 颜色空间理论, 将 2 幅输入图像转换到相关

的 $l\alpha\beta$ 空间, 再进行点对点的加计算, 改善了仿真结果图像的视觉效果。本文以较为常见的几种烙画作为实验对象, 利用计算机来模拟其表面色彩及纹理效果, 让这一艺术形式能在计算机上得到重现。

2 基于轮选择粒子群优化的烙画学习

2.1 粒子群优化算法基本思想

粒子群优化算法是一种仿生进化算法, 属于全局优化算法^[5]。其主要优点在于算法易于编程实现、收敛性保证以及收敛速度快等。该算法可以简要描述为: 设目标搜索空间为 D 维, 粒子群总数为 m , 第 i 个粒子在搜索空间的位置用向量 $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})^T$ 表示, 该粒子搜索到的最优位置为 $P_i = [P_{i1}, P_{i2}, \dots, P_{iD}]^T$, 整个粒子群搜索到的最优位置为 $P_g = [P_{g1}, P_{g2}, \dots, P_{gD}]^T$, 在迭代过程中, 第 i 个粒子以速度 $V_i = [V_{i1}, V_{i2}, \dots, V_{iD}]^T$ 在搜索空间飞行。则单个粒子的更新其速度和位置公式可以表示为:

$$v_{id}^{k+1} = w^k v_{id}^k + c_1 r_1 (p_{id} - x_{id}^k) + c_2 r_2 (g_{id} - x_{id}^k) \quad (1)$$

基金项目: 陕西省 2009 年“13115”科技创新工程基金资助项目(2009ZDGC-06)

作者简介: 王 东(1968 -), 男, 讲师、博士研究生, 主研方向: 数字图像处理, 色彩管理; 周世生, 教授、博士、博士生导师

收稿日期: 2010-06-08 **E-mail:** wd9956@sina.com

$$x_{id}^{k+1} = x_{id}^k + w^k v_{id}^k \quad (2)$$

其中, k 为迭代次数; c_1 、 c_2 为学习因子; r_1 、 r_2 表示(0, 1)之间的随机数; w^k 为惯性权重, 进化过程为:

$$w^k = w_{\max} - \frac{w_{\max} - w_{\min}}{k_{\max}} k \quad (3)$$

其中, k_{\max} 为最大进化代数; w_{\min} 和 w_{\max} 为 w^k 的取值范围。

上述算法即为 PSO 最原始的理论基础, 与其他进化算法一样, PSO 算法也面临着粒子收敛过快, 导致早熟的问题, 使算法丧失进一步搜索的能力。这主要是由于 PSO 算法是通过粒子跟踪 2 个“极值”从而达到寻优目的。在寻优过程中, 如果整个种群的最优位置处于一个局部极值位置, 则所有粒子将迅速向 p_g 靠拢, 从而导致 PSO 算法过早收敛到一个局部极值点。

2.2 粒子群优化算法的改进

粒子群优化实质上是一种基于块匹配的搜索算法。Hertzmann 在其算法中使用了最佳匹配像素搜索算法, 该算法所要达到的目标是要在高维空间中找到最精确的点, 但对于风格学习而言, 要追求的是视觉上的相似, 所以能够接受的解可以不是最优解, 这样便可以使算法在实现加速的同时, 避免陷入局部最优而停滞。

由上述的分析可知, PSO 算法容易陷入局部最优的一个重要原因是当 p_g 处于一个局部极值位置后, p_g 将不再变化, 而且所有粒子将迅速向其靠拢。为解决这个问题, 本文改进了选取 p_g 的策略, 不是简单地选取适应度最大的粒子为 p_g , 而是采用轮选择的方法, 从若干个适应度最大的粒子中选取一个为 p_g 。具体做法是: 设整个粒子群中前 L 个最优位置为 $P_g^1, P_g^2, \dots, P_g^L$, 其对应的适应度分别为 g^1, g^2, \dots, g^L , 按轮选择方法, $P_g^i (i=1, 2, \dots, L)$ 被选为 p_g 的概率为:

$$P_i = \frac{f_i}{\sum_{j=1}^L f_j}, \quad \sum_{i=1}^L P_i = 1 \text{ 且 } P_i \geq 0 \quad (4)$$

在上述算法中, 函数 $\text{random}()$ 产生一个(0, 1)区间内的随机数, $\text{fitness} = \sum_{j=1}^n \text{fitness}[j]$ 。通过比较出随机数 rand 与适应值 fitness 的前 j 项部分和之间的大小关系, 确定随机数 rand 落入的区间。这样, 算法不仅能够保留基本 PSO 算法的特点, 即通过跟踪 2 个“极值”而达到寻优的目的, 而且当整个粒子群的最优位置处于局部最优时, 可以通过选择不同的 p_g 使粒子群朝其他方向移动, 以避免所有粒子迅速朝该局部最优位置靠拢, 从而降低算法过早收敛于局部最优解的几率。尽管此时选中的粒子不是目前的最优粒子, 但也是粒子群中的次优粒子, 从而能够保证整个粒子群还是朝好的方向飞行。

2.3 图像学习的粒子群加速算法

2.3.1 粒子属性的定义

给定 3 幅源图像分别为 A 、 A' 和 B , 设 B' 为待合成图像, 其对应的特征向量以 $A(p)$ 、 $A'(p)$ 、 $B(q)$ 、 $B'(q)$ 表示。在算法处理中, 需要把源图 A' 中的像素点 p 的特征向量赋值给目标图像 B' 的像素 q 点, 为此建立一个以 p 为索引的数据结构 $s(\cdot)$, 即 $s(p) = q$ 。在源图像 A 中的第 i 个粒子对应像素点的位置, 坐标记为 $(\text{Present}X_i, \text{Present}Y_i)$, 粒子 i 经历的最好位置的坐标记为 $(\text{Lbest}X_i, \text{Lbest}Y_i)$, 粒子在该位置取得的适应值记为 Lbest_i , 显然有 $Gbest = \min\{\text{Lbest}_1, \text{Lbest}_2, \dots, \text{Lbest}_n\}$ 。

每个随机粒子的初始运动向量 $V_i = (V_{ix}, V_{iy})$, 另外还需定义 2 个运动向量 $L_i(V_{ix}, V_{iy})$ 、 $G_i(V_{ix}, V_{iy})$, 这样每个粒子便可以据此比较当前最优解和全局最优解, 以改变其运动方向和速度。粒子在运动过程中的速度和位置更新可由式(1)和式(2)来确定。

2.3.2 粒子适应值及相似阈值的选取

对于搜索空间中的每个粒子, 在搜索过程中任一时刻都要计算其适应值。适应值表示当前粒子所在位置的 L 邻域与待合成像素 L 邻域的匹配程度。定义 D 为待合成像素的 L 邻域, S 为粒子所在位置的 L 邻域。则粒子适应值的计算公式如下:

$$D(p, q) = \sum_{i=0, j=0}^{4, 4} \text{sqr}t \left\{ \left[(q(i, j)_R - p(i, j)_R)^2 + (q(i, j)_G - p(i, j)_G)^2 + (q(i, j)_B - p(i, j)_B)^2 \right] \right\} \quad (5)$$

其中, $q(i, j)_R$ 、 $q(i, j)_G$ 、 $q(i, j)_B$ 和 $p(i, j)_R$ 、 $p(i, j)_G$ 、 $p(i, j)_B$ 分别为目标图像和源图像的 L 邻域内 (i, j) 处的 R 、 G 、 B 颜色分量。

相似性匹配阈值表示目标图像 B 中的任一像素点 q 与源图像中与其相似度最高的像素点 p 的匹配程度, 即为两像素点之间的特征向量的距离 D_{\max} 。通过逐一比较 2 个邻域中像素点的 RGB 值得到累积误差, 累积误差越小表示相似度越高。相似性匹配阈值同时用以确定粒子群优化算法的搜索终止门限。若某一时刻粒子群的最优适应值低于 D_{\max} , 则可中止迭代过程, 并将取得最优适应值的粒子所在像素点作为最优解返回。相似性阈值的计算公式如下:

$$d_{\max} = \varepsilon \left\{ \text{sqr}t \left[\sum_{i=1}^k R(p_{\text{out}}^i)^2 + G(p_{\text{out}}^i)^2 + B(p_{\text{out}}^i)^2 \right] \right\} \quad (6)$$

其中, $R()$ 、 $G()$ 、 $B()$ 分别表示像素点的三原色值; k 为 L 邻域内像素点个数; ε 为误差系数, 是个经验值。对于随即性较强的烙画媒质纹理, 如毛毡、宣纸等, ε 值要选择小些; 而对于比较规整或连续性较强的媒质纹理, ε 值应选择大些。

2.3.3 算法描述

首先随机初始化源图像 A 、 A' 和 B , 并计算出个像素的特征向量; 按照行扫描顺序从左至右依次选取图像 B 中的像素点 q , 计算适应值 D_i ; 设定粒子 i 当前的最优位置为 $P_i = X_i$, 所对应的适应值为 $D_{\text{best}i} = D_i$, 在源图像 A 中搜索出适应度最大的 L 个粒子, 其位置及适应值分别为 $P_g^1, P_g^2, \dots, P_g^L$ 和 g^1, g^2, \dots, g^L , 且满足 $g^1 \leq g^2 \leq \dots \leq g^L$, 应用式(4)计算此 L 个粒子成为全局最优位置的概率, 根据所产生的随机数 $\zeta \in (0, 1)$, 与各概率值进行比较, 确定全局最优解。继而将源图像 A' 中的像素点 P_i 的特征向量赋值给目标图像 B' 中对应的像素点 q_i , 并进行下一像素点的合成处理; 根据式(1)、式(2)更新其余粒子的位置 p_{best} 和速度 v 及适应值; 若 $D_i > D_{\text{best}i}$, 则 $P_i = X_i$ 、 $D_{\text{best}i} = D_i$; 重复上述步骤, 直至达到最大的迭代次数, 目标图像 B' 中所有像素点合成完毕。在实际应用中, 一般取全部解向量的 20% 为宜^[6]。

2.4 图像的融合增强

经过上述优化算法所得到的图像风格化学学习结果, 通常不能满足诸如木板等自然纹理的视觉效果, 因此, 需要人为进行视觉效果增强。笔者在实验中发现, 根据 lab 颜色空间的理论, 将 2 幅输入图像转换到相关的 lab 空间, 经 lms 空间及其对数转换, 再进行点对点的加计算, 最后再返回到 RGB 空间, 图像还原后可以避免图像像素值溢出。设 $F(x, y)$

和 $G(x, y)$ 是 2 张不相关的图像, 进行加运算后得到新的图像 $S(x, y)$, 加运算的数学表示为:

$$S(x, y) = F(x, y) + G(x, y) \quad (7)$$

算法过程如下:

(1) 将 $F(x, y)$ 和 $G(x, y)$ 图像的像素信息从 RGB 颜色空间变换至 lms 颜色空间。

(2) 用 $L = \lg l$ 、 $M = \lg m$ 、 $S = \lg s$ 来代替 lms 的值。

(3) 运用主元分析理论, 计算出 LMS 和 $l\alpha\beta$ 之间的变换矩阵, 并利用这个变换矩阵把 LMS 基变换到一组正交基, 把 RGB 颜色空间转换到一个正交空间 $l\alpha\beta$, 消除这些空间的各个分量之间所具有的强相关性。

(4) $F(x, y)$ 和 $G(x, y)$ 转移到 $l\alpha\beta$ 颜色空间下后, 得到新的图像 $F'(x, y)$ 和 $G'(x, y)$, 然后再进行加运算后得到新的图像 $S'(x, y)$, $S'(x, y) = F'(x, y) + G'(x, y)$ 。

(5) 在 $l\alpha\beta$ 颜色空间下完成图像的加运算后, 需要把这个新的 $l\alpha\beta$ 值通过逆变换转变成 RGB 值, 从而产生需要的新结果图像。

3 实验与分析

本次实验以 VC++6.0 为开发平台, 在 P4 3.4 GHz、内存 1 GB 的计算机上实现。与 Hertzmann 的 ANN 算法的运行时间进行比较, 图 1~图 3 为以花鸟、山水及人物为学习对象的结果。

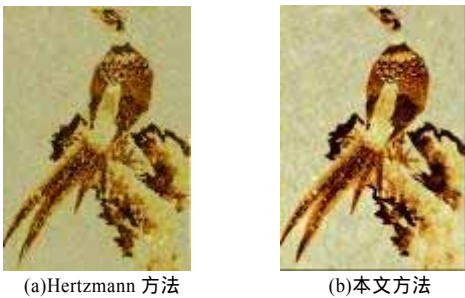


图 1 毛毡烙画学习结果

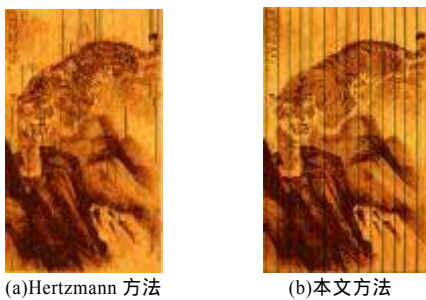


图 2 木板条烙画学习结果



图 3 木板烙画学习结果

从毛毡烙画学习效果图中可以看出, Hertzmann 的 ANN 算法的运行结果与本文基于轮选择的 PSO 算法所获得的效果相差不大, 这是由于毛毡纹理随机性较大, 人眼不易分辨其接缝的缘故。在木板条烙画图像的学习效果中, 由于木板条这种整齐的纹理排列, 无论是 Hertzmann 还是本文的改进 PSO 算法, 均显得力不从心, 这一点从木板条烙画图像学习效果所显示的效果中可以得到验证。因此, 本文运用了 3 种所提出的图像融合增强方法, 作进一步的纹理修饰。在木板烙画的模拟中, 本文选择了菲律宾木三合板作为承载的媒质, 是一种通过学习直接传输的方式。可以看出, Hertzmann 的 ANN 算法在头像的细节刻画方面要好些, 而周围的木纹线的比较细碎, 而本文所采用的方法在木纹的连续性方面要优于 Hertzmann 算法。

Hertzmann 的 ANN 算法与本文的轮选择 PSO 算法时间对比如表 1 所示。在 3 种类型的烙画效果中, 源图像及目标图像尺寸分别为: 285×319, 283×400; 283×527, 250×498; 197×263, 197×257。可以看出, 本文算法较 Hertzmann 算法有一定速度提升。

表 1 本文算法与 Hertzmann 算法耗时比较

烙画类型	本文算法耗时/s	Hertzmann 算法耗时/s	加速倍数
毛毡	237	1559	6.58
木板条	234	1432	6.12
三合板	275	1741	6.31

4 结束语

图像艺术风格处理在非真实感图形学中一直是人们关注的热点, 许多学者尝试用多种方法实现精美的艺术效果。数字烙画艺术可广泛应用于装饰设计、电子贺卡、广告以及游戏、动画制作等领域。本文在烙画风格的模拟过程中取得了一定的阶段性成果。但本文算法还存在着许多需要改进的地方。本文所实现的烙画效果, 尤其是在自然纹理的学习方面, 与真实的木纹还有一定的差距。烙画的承载媒质很多, 如何一步研究具有一般性、能够适应更多烙画风格类型的方法是后续研究的关键所在。

参考文献

- [1] 王景波, 陆 玲. 基于 OpenGL 的 3D 果实造型研究[J]. 计算机工程, 2010, 36(4): 279-280.
- [2] Hertzmann A, Jacobs C E, Oliver N, et al. Image Analogies[C]//Proc. of SIGGRAPH'01. Los Angeles, California, USA: [s. n.], 2001.
- [3] Mount D M. ANN Programming Manual[D]. College Park, Maryland, USA: University of Maryland, 2006.
- [4] Ruderman D L, Cronin T W, Chiao C C. Statistics of Cone Responses to Natural Images: Implications for Visual Coding[J]. Journal of the Optical Society of America, 1998, 15(8): 2036-2045.
- [5] Kennedy J, Eberhart R C. Particle Swarm Optimization[C]//Proc. of IEEE International Conf. on Neural Networks. Perth, Australia: [s. n.], 1995.
- [6] Premalatha K. A New Approach for Data Clustering Based on PSO with Local Search[J]. Computer and Information Science, 2008, 1(4): 139-145.

编辑 顾姣健