

# 一种新的图像合成方法

刘则毅<sup>1),2)</sup> 徐高奎<sup>2)</sup> 丁雅斌<sup>3)</sup>

<sup>1)</sup>(深圳大学理学院, 深圳 518060) <sup>2)</sup>(天津大学理学院数学系, 天津 300072)

<sup>3)</sup>(天津大学精密测试技术及仪器国家重点实验室, 天津 300072)

**摘要** 提出了一种新的图像分割与合成方法。对于图像的分割,提出了新的切割图法,先是手工把一些明显属于物体和明显属于背景的像素分割出来,然后设置了新的能量函数,求它取最小值时的分割结果。对于图像的合成,提出了梯度比较法,在保留物体基本颜色特征的同时,改变了物体的光亮度,使其与新背景的光亮度一致,从而合成图像的光照效果更真实。

**关键词** 图像分割 图像合成 切割图法 梯度比较法

**中图分类号**: TP391 **文献标识码**: A **文章编号**: 1006-8961(2007)03-0432-06

## A New Method for Image Composition

LIU Ze-yi<sup>1),2)</sup>, XU Gao-kui<sup>2)</sup>, DING Ya-bin<sup>3)</sup>

<sup>1)</sup>(Science College of Shenzhen University, Shenzhen 518060)

<sup>2)</sup>(Department of Mathematics of Science College, Tianjin University, Tianjin 300072)

<sup>3)</sup>(State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072)

**Abstract** This paper presents a new image segmentation and composition method. This method first separates the pixels which obviously belong to objects or background, then gets the cutout whose corresponding value of the new energy function is the minimum. Afterwards the cutout is fused into the target image using gradient comparing method. The gradient comparing method ensures the compliance of illumination in the resultant image meanwhile preserves the color characteristic of the object. It can be easily concluded from the experiments that the new composition method has extended applications and the resultant image has a better lighting than those produced by other methods.

**Keywords** image segmentation, image composition, graph cut method, gradient comparing method

## 1 引言

图像的分割与合成是指把一张图像中感兴趣的物体分割出来,合成到另一张图像中,这在图像编辑中有十分广泛的应用,特别是对于一些现实中很困难或根本不可能实现的场景尤为重要,因为这种场景实际不存在,它的图像无法通过实地拍摄获得,但可以通过图像合成来达到这种效果,如科幻电影与动画片的制作。近年来,国内外的学者已经对图像

的分割与合成进行广泛的研究,提出了很多实现的方法,下面就对这些方法进行简单介绍。

### 1.1 已有方法

1.1.1 把感兴趣的物体从源图像中准确地分割出来,主要有3类方法:

(1) 轮廓线逼近法(Active Contours)或蛇线法(Snakes)<sup>[1]</sup> 首先根据图像的梯度信息与曲线的特征信息建立一个能量函数,然后逐步移动初始曲线,直至它所对应的能量函数值达到最小,此时的闭曲线即认为是物体的轮廓线。缺点是自动性不强,结

**基金项目**:国家自然科学基金项目(60275012);广东省普通高校自然科学研究重点项目(04Z010);广东省自然科学基金项目(031804);深圳市科技计划项目(200341)

**收稿日期**:2005-04-13;**改回日期**:2005-11-28

**第一作者简介**:刘则毅(1958~),男,教授。1997年于天津大学系统工程研究所获系统工程专业博士学位。主要研究方向为计算机图形学和智能计算方法。E-mail: zeyiliu@szu.edu.cn

果不够准确。

(2) 智能裁减法 (Intelligent Scissors)<sup>[2]</sup> 先在物体内部找几个种子点,然后人工交互在物体的轮廓线附近选择控制点,通过计算已知种子点到人工控制点的最优路径来求得物体的轮廓线。缺点仍是自动性不强,结果不够准确,而且还很费时。

(3) 切割图法 (Graph Cuts)<sup>[3-5]</sup> 先建立物体和背景的颜色模型,也就是物体和背景颜色的分布情况,然后根据这个信息及图像的颜色信息建立特殊形式的能量函数,最后用文献[6]、[7]中的方法求能量函数取最小值时的分割结果。此类方法自动性强、速度快、准确性高,是现在比较流行的一类方法。

1.1.2 把分割出来的物体合成到目标图像中,也大致有3类方法:

(1) 直接复制法<sup>[2,5]</sup> 不改变体的颜色,光照,最多只对其方向大小进行变化。此方法简单但是当物体与新背景的光照条件相差很大时,得到的图像不真实。

(2) Poisson 合成法<sup>[8,9]</sup> 它是在 RGB 向量每个方向上用 Poisson 方程对物体与新背景进行合成,这在一定程度上可以改善合成图像的光照效果,但是因为物体的颜色随着背景颜色的变化发生改变,所以一旦源图像与目标图像的背景颜色相差很大,合成图像中物体的颜色就会严重失真。

(3) 3 维光照法<sup>[10]</sup> 先从多张图像重建出物体与新场景的 3 维结构,然后在 3 维的情况下详细计算物体各个部分的实际光照。此方法无疑准确度是最高的,得到的合成结果也最好,但是需要的条件太多,而且实现起来很复杂,这对于一般的要求不太高的图像合成来说是不实际也是没有必要的。

## 1.2 本文的方法

本文提出的切割图法先把图像中许多明显属于物体和明显属于背景的像素手工分割出来,根据这些像素分别建立起物体与背景的颜色模型,然后构造了一个新的能量函数,并在剩余的像素上求此能量函数取最小值时的分割结果,即得到分割出来的物体。

对于图像合成,本文提出了梯度比较法,它首先计算合成图像中物体各像素在新背景光照下的灰度值,然后令源图像中物体各像素的 RGB 向量随着该像素在合成图像中灰度值的变化进行相应变化,从而得到合成图像中物体各像素的 RGB 向量。

## 2 主要算法

### 2.1 把感兴趣的物体从源图像中分割出来

#### 2.1.1 切割图法的一般步骤

先建立物体和背景的颜色模型,也就是物体和背景颜色的分布情况,然后决定图像中每个像素是属于物体还是属于背景。一般来说,判断一个像素属于物体还是背景,需要考虑两点,一是这个像素本身的颜色,如果相对于背景颜色来说,它的颜色与物体的颜色更接近,那么它属于物体的可能性大于属于背景的可能性。二是它周围的像素属于哪部分,如果它周围的像素都属于物体,那么它也有可能属于物体,而且它的颜色跟周围像素的颜色越相似,这个可能性越大。把这两个因素进行综合,建立能量函数:

$$E(X) = \sum_{p \in P} E_1(p, x_p) + \lambda \sum_{(p, q) \in N} E_2(p, q, x_p, x_q) \quad (1)$$

其中,  $E_1$  体现了上面的第 1 个因素,即像素本身的颜色在决定它的归属时的作用,  $E_2$  体现了第 2 个因素,即一个像素周围像素的归属情况对它的归属的影响。  $P$  是图像中所有像素的集合,  $p \in P$  表示  $p$  是图像中的像素。  $N$  为图像中所有相邻像素对的集合,  $(p, q) \in N$  表示  $p, q$  在图像中是相邻像素。  $X$  为图像的一种分割结果,也就是对每个像素指出它是属于物体还是属于背景,表示为  $X = \{x_p | p \in P\}$ , 其中  $x_p$  为像素  $p$  的标志,它的取值只有两种情况:

$$x_p = \begin{cases} 1 & p \text{ 在分割结果 } X \text{ 中属于物体} \\ 0 & p \text{ 在分割结果 } X \text{ 中属于背景} \end{cases}$$

$E(X)$  为分割结果  $X$  对应的能量函数值。  $E_1$  称为似然函数,  $E_1(p, x_p) > 0$  是像素  $p$  的标志为  $x_p$  时的惩罚值,它的颜色跟物体(背景)的颜色越相似,  $E_1(p, x_p = 1) (E_1(p, x_p = 0))$  值越小;  $E_2$  称为优先权函数,  $E_2(p, q, x_p, x_q)$  是相邻像素  $p, q$  的标志分别为  $x_p, x_q$  时的惩罚值。如果  $x_p = x_q, E_2(p, q, x_p, x_q) = 0$ ; 如果  $x_p \neq x_q, E_2(p, q, x_p, x_q) > 0$ , 并且像素  $p, q$  的颜色越相似,  $E_2(p, q, x_p, x_q)$  越大。  $\lambda$  反映似然函数  $E_1$  与优先权函数  $E_2$  在总的能量函数  $E$  中的相对比重。最后用文献[6]、[7]中的方法求此能量函数取最小值时的分割结果。

#### 2.1.2 本文切割图法的具体操作

要把物体从图像中分割出来,也就是对图像中每个像素判断出它是属于物体还是属于背景。图像

中有些部分很明显属于物体,如图 1(a)、(c)中标  $F$  的部分,可以用曲线把它们与其他部分分开。同样有些部分很明显属于背景,如图 1(a)、(c)中标  $B$  的部分,也可以用曲线把它们分开。对于其余部分,



(a) 待分割的图像

(b) 图像(a)的分割结果

(c) 待分割的图像

(d) 图像(c)的分割结果

图 1 图像分割

Fig. 1 Image segmentation

### 步骤 1 建立物体与背景的颜色模型

用 Orchard<sup>[11]</sup> 中的方法把  $F$  中的像素分为  $m$  块,  $m$  既可以根据具体情形提前指定,也可以通过设定的条件由算法自己产生。对于每一块,计算其中所有像素 RGB 向量的平均值及其协方差矩阵,如第  $i$  块  $P_i$  (像素个数为  $I$ ) 的平均值与协方差矩阵为

$$\bar{C}_i = \frac{1}{I} \sum_{p \in P_i} C_p \quad (2)$$

$$L_i = \frac{1}{I} \sum_{p \in P_i} (C_p - \bar{C}_i)(C_p - \bar{C}_i)^T \quad (3)$$

其中,  $C_p$  为像素  $p$  的 RGB 列向量。从而得到物体的颜色模型:  $m$  个三元正态分布函数  $\{N(\mu_1, \Sigma_1), N(\mu_2, \Sigma_2), \dots, N(\mu_m, \Sigma_m)\}$ , 其中  $\mu_i = \bar{C}_i, \Sigma_i = L_i, i = 1, 2, \dots, m$ 。

同样方法得到背景的颜色模型:  $n$  个三元正态分布函数  $\{N(\mu_1^*, \Sigma_1^*), N(\mu_2^*, \Sigma_2^*), \dots, N(\mu_n^*, \Sigma_n^*)\}$ 。

### 步骤 2 建立能量函数

对于任意一个像素  $p \in U$ , 它的 RGB 向量在物体颜色模型的第  $i$  个正态分布中的概率为

$$\frac{1}{(2\pi)^{\frac{3}{2}} \det(\Sigma_i)^{\frac{1}{2}}} \exp(- (C_p - \mu_i)^T \Sigma_i^{-1} (C_p - \mu_i)) \quad (4)$$

选择这  $m$  个概率中最大的概率作为像素  $p$  的颜色与物体颜色的相似程度, 记为

$$K_p^F = \max_{1 \leq i \leq m} \frac{1}{(2\pi)^{\frac{3}{2}} \det(\Sigma_i)^{\frac{1}{2}}} \exp(- (C_p - \mu_i)^T \Sigma_i^{-1} (C_p - \mu_i)) \quad (5)$$

同理像素  $p$  的颜色与背景颜色的相似程度为

如图 1(a)、(c) 中标  $U$  的部分, 有的像素属于物体, 有的属于背景, 所以要把物体从图像中分割出来, 也就是要对  $U$  中每个像素判断它是属于物体还是属于背景。

$$K_p^B = \max_{1 \leq i \leq n} \frac{1}{(2\pi)^{\frac{3}{2}} \det(\Sigma_i^*)^{\frac{1}{2}}} \exp(- (C_p - \mu_i^*)^T \Sigma_i^{*-1} (C_p - \mu_i^*)) \quad (6)$$

将  $E_1$  定义为

$$E_1(p, x_p = 1) = \frac{\log K_p^F}{\log K_p^F + \log K_p^B} \quad (7)$$

$$E_1(p, x_p = 0) = \frac{\log K_p^B}{\log K_p^F + \log K_p^B} \quad (8)$$

可见像素  $p$  的颜色与物体的颜色越相似,  $K_p^F$  越大,  $\log K_p^F$  越大, 那么  $E_1(p, x_p = 1) = \frac{\log K_p^F}{\log K_p^F + \log K_p^B}$  越小, 从而在最后得到的分割结果中,  $x_p$  越有可能取到 1, 即像素  $p$  越有可能属于物体。像素  $p$  的颜色与背景颜色的关系同理。

因为  $E_2$  定义在相邻像素对上, 所以首先需要到像素的相邻关系进行定义。本文的像素相邻关系这么定义, 每个像素只与它的上下左右 4 个像素相邻, 如图 2 所示, 中间标 0 的像素只与 4 个标 1 的像素相邻。

$E_2$  的意义是对相邻像素在分割结果中被分开 (分割结果中相邻的两个像素分别属于物体与背景) 的情况进行惩罚。一般来说如果相邻像素颜色越相似, 那么它们越有可能同属物体或同属背景, 从而在分割结果中越不应该被分开, 所以如果这时  $x_p \neq x_q$ , 那么  $E_2$  值应该越大, 反之越小。故  $E_2$  定义为

$$E_2(x_p, x_q) = |x_p - x_q| \cdot g(C_p, C_q) \quad (9)$$

其中,  $|x_p - x_q|$  是  $x_p$  与  $x_q$  之差的绝对值,  $g(\xi) =$

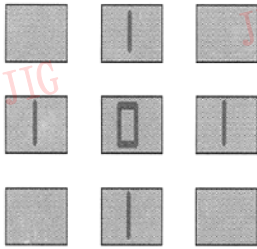


图2 相邻像素

Fig.2 Neighborhood pixels

$\frac{1}{1+\xi}, C_{p,q} = \|C_p - C_q\|^2, C_p, C_q$  分别为像素  $p, q$  的 RGB 列向量,  $\|C_p - C_q\|^2$  是  $C_p - C_q$  的模的平方。

步骤3 求解

用文献[6]、[7]中的方法求此能量函数取最小值时的分割结果(如图1)。

2.2 把分割出来的物体合成到目标图像中

每个图像都可以看作是区域  $S \in \mathbf{R}^2$  (如图3, 其中  $S$  的大小和形状与图像的大小和形状相同) 上的3维向量函数, 函数在每个像素上的值就是图像在该像素上的 RGB 向量。设  $f_1, f_2$  分别是源图像与目标图像, 从  $f_1$  中分割出来的物体所覆盖的区域为  $\Omega (\Omega \in S, \text{如图3})$ , 现在要考虑的就是如何把图像  $f_1$  在  $\Omega$  中的部分与目标图像  $f_2$  在  $\Omega$  以外的部分二者合成得到新图像  $g$ , 并使  $g$  看起来真实自然。

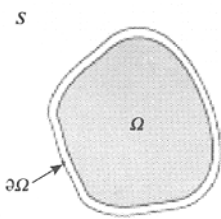


图3 S与Omega的关系

Fig.3 Relation between S and Omega

最简单的方法就是把  $f_1$  中属于  $\Omega$  的部分直接复制到  $f_2$  中相应的区域, 这种方法实现起来容易速度快, 而且对于不少的图像合成总体效果也不错, 事实上很多图像合成<sup>[2,5]</sup> 也就是这样做的, 但是这种方法对于下面的情况就无法得到满意的结果。当图像  $f_1, f_2$  拍摄时光照条件相差很大, 其中一幅图像很亮, 另一幅图像很暗, 如果直接把  $f_1$  中的  $\Omega$  部分复制到  $f_2$  中, 就会得到物体很暗但背景很亮, 或者

物体很亮但背景很暗的合成图像, 如图4(b), 图5(b)。显然这种图像是不真实的, 因为现实中如果光照强, 那么物体与背景都应该亮; 光照弱, 两者都应该暗。所以要想得到光照合理的图像, 必须对源图像中物体的亮度做出调整, 使其与目标图像的亮度一致, 下面是本文提出的调整方法。



(a) 目标图像

(b) 直接复制法的结果



(c) Poisson法的结果

(d) 梯度比较法的结果

图4 把图3(a)中的物体与新背景合成

Fig.4 Composition of object in Fig.3(a) with new background

先求出图像  $f_1, f_2$  的灰度图像  $\hat{f}_1, \hat{f}_2$ , 然后把满足下面条件的  $\hat{g}$  作为最后合成图像  $g$  的灰度图像:

$$\min \sum_{\langle p, q \rangle \cap \Omega \neq \emptyset} [(\hat{g}(p) - \hat{g}(q)) - (\hat{f}_1(p) - \hat{f}_1(q))]^2 \tag{10}$$

$$\hat{g}|_{S/\Omega} = \hat{f}_2|_{S/\Omega} \tag{11}$$

其中,  $\langle p, q \rangle$  表示像素  $p, q$  是相邻像素,  $\langle p, q \rangle \cap \Omega \neq \emptyset$  说明像素  $p, q$  中至少有一个属于  $\Omega$ 。  $S/\Omega$  指区域  $S$  中除去区域  $\Omega$  外的部分,  $\hat{g}|_{S/\Omega} = \hat{f}_2|_{S/\Omega}$  是指  $\hat{g}$  与  $\hat{f}_2$  在  $S/\Omega$  中每个像素处的值都相同。式(10)使得合成图像中物体相邻像素的相对亮度与在源图像中的相对亮度基本相同, 式(10)与式(11)使得合成图像中物体与背景的亮度一致。



图 5 把图 3(c)中的物体与新背景合成

Fig. 5 Composition of object in Fig. 3(c) with new background

由上式得到合成图像中每个像素的灰度之后,再如下计算合成图像中每个像素的 RGB 向量

$$g(\mathbf{p}) = \begin{cases} f_2(\mathbf{p}) & \mathbf{p} \in S/\Omega \\ \hat{g}(\mathbf{p}) \cdot f_1(\mathbf{p}) & \mathbf{p} \in \Omega \end{cases} \quad (12)$$

式(12)使得合成图像中物体在每个像素处的 RGB 相对值与在源图像中相同,又由于物体相邻像素的相对亮度基本不变,从而物体颜色的基本特征并未改变,只是整体亮度发生变化,并且变化后的亮度与背景亮度一致,合成图像看起来真实自然,如图 4(d),图 5(d)。

### 3 实例及分析

图 4 是图像合成的一个例子,其中图 4(a)是目标图像,图 4(b)是把图 3(a)中分割出来的人物直接复制到目标图像中得到的结果,图 4(c)是用 Poisson 合成法得到的结果,图 4(d)是用梯度比较法合成的结果。显然图 4(b)的结果不合实情,因为被太阳直接照射的部分与被遮挡住的部分亮度一样。图 4(c)中的整体亮度比较合理,但是物体颜色受到背景颜色的影响变得发黄,偏离了物体本来的

颜色。图 4(d)中物体的颜色基本未变,只是亮度发生变化,被太阳直接照射的部分变得更亮,被遮挡住的部分变得更暗,比较真实。图 5 是另一个图像合成的例子,图 5(a)仍然是目标图像,图 5(b)是把图 3(c)中分割出来的头像直接复制到图 5(a)中得到的结果,图 5(c)是 Poisson 合成法的结果,图 5(d)是用梯度比较法得到的结果。图 5(b)中背景很暗,头像很亮,显然不合理。图 5(c)中背景与头像都很暗,但是由于背景颜色中蓝色的比例比较大,使得头像也整体变蓝,不符合事实。图 5(d)中头像的颜色特征基本不变,亮度随着背景变暗也相应的变暗,符合实际情况。可见在上面两个例子中直接复制法与 Poisson 融合法都得不到令人满意的结果,因为直接复制法要求源图像与目标图像亮度差不多, Poisson 合成法要求源图像与目标图像中背景颜色比较相似,而本例中这两个条件都不满足。如果采用 3 维光照计算,仅从这两幅图像并不能得到物体与背景的 3 维结构,故此法也不适用,而且即使得到它们的 3 维结构,计算也相当复杂,对于像这种要求并不高的情况显得没有必要。相比较而言本文方法并不需要太复杂的计算,就使得合成图像中亮度统一,并且保持了物体的颜色特征,使得结果图像真实可信,所以是一个比较实用的方法。

### 4 结论

本文提出的切割图法,在建立物体与背景的颜色模型时使用了大量的像素,并且计算时采用了新的能量函数,所以得到的分割结果准确,又由于计算能量函数的最小值时只需考虑  $U$  中的像素,而不是整个图像的所有像素,因而速度快,但这个方法的缺点是人工交互多,开始时需要手工把它分为 3 部分,不过这对于现在的图形工具来说很容易做到。对于图像合成,梯度比较合成法既保证物体的颜色特征基本不变,又使得合成图像中物体与背景的亮度一致,克服了直接复制法与 Poisson 合成法的不足。实验证明它的应用范围更广,得到的合成图像效果更好。

### 参考文献 (References)

- 1 Kass M, Witkin A, Terzopoulos D. Snakes: active contour models [A]. In: Proceedings of the First International Conference on Computer Vision[C], London, England, 1987: 259 ~ 68.

- 2 Mortensen F N, Barrett W A. Intelligent scissors for image composition[A]. In: Proceedings of the ACM SIGGRAPH: 22nd International Conference on Computer Graphics and Interactive Techniques[C], Los Angeles, CA, 1995: 191 ~ 198.
- 3 Boykov Y, Jolly M P. Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images[A]. In: International Conference on Computer Vision [C], Vancouver, B C, Canada, 2001: 105 ~ 112.
- 4 Rother C, Blake A, Kolmogorov V. Grabcut-interactive foreground extraction using iterated graph cuts [A]. In: Proceedings of the SIGGRAPH Conference[C], New York, USA, 2004: 309 ~ 314.
- 5 Li Yin, Sun Jian, Tang Chi-keung, *et al.* Lazy snapping[A]. In: Proceedings of the SIGGRAPH Conference [C], New York, USA, 2004: 303 ~ 308.
- 6 Boykov Y, Veksle O, Zabih R. Fast approximate energy minimization via graph cuts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(11): 1222 ~ 1239.
- 7 Boykov Y, Kolmogorov V. An experimental comparison of mincut/max-flow algorithms for energy minimization in vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(9): 1124 ~ 1137.
- 8 P'erez P, Gangnet M, Blake A. Poisson image editing[A]. In: Proceedings of the SIGGRAPH Conference [C], New York, USA, 2003: 313 ~ 318.
- 9 Agarwala A, Dontcheva M, Agrawala M, *et al.* Interactive digital photomontage[J]. ACM Transactions on Graphics, 2004, 23(3): 294 ~ 302.
- 10 McNerney P. A stereo-based approach to digital image compositing [D]. Boston MA, USA: Department of Electrical and Computer Engineering, Boston University, 2003.
- 11 Orchard M, Bouman C. Color quantization of images [J]. IEEE Transactions on Signal Processing, 1991, 39(12): 2677 ~ 2690.