

一种基于时空一致性的视频抠取方法^{*}

丁子昂, 夏佳志, 管宇, 陈为⁺, 彭群生

(浙江大学 CAD&CG 国家重点实验室, 浙江 杭州 310058)

Interactive Space-Time Video Matting Approach

DING Zi-Ang, XIA Jia-Zhi, GUAN Yu, CHEN Wei⁺, PENG Qun-Sheng

(States Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310058, China)

+ Corresponding author: E-mail: chenwei@cad.zju.edu.cn, http://www.cad.zju.edu.cn

Ding ZA, Xia JZ, Guan Y, Chen W, Peng QS. Interactive space-time video matting approach. Journal of Software, 2008,19(Suppl.):10-18. <http://www.jos.org.cn/1000-9825/19/s10.htm>

Abstract: This paper presents an interactive matting approach for efficiently extracting alpha mattes and foreground objects from video sequences. Beginning from user-specified strokes across space and time, the paper formulates their expansions in the video volume as a Laplacian equation, resulting in a coarse alpha matte. It then employs a novel spacetime alpha matting technique that makes use of local statistics and neighboring information and converges to a globally optimal alpha matte in few iterations. Finally, the paper derives a new global cost function to reconstruct the foreground color of the whole video volume, which faithfully preserves the spatio-temporal coherence. The computation in each step can be reformulated as solving a set of linear equations, allowing users to quickly extract high-quality alpha matte and foreground objects, even for data sets with ten million pixels. Experimental results on complex natural video sequences demonstrate the high quality and efficiency of the proposed approach.

Key words: video matting; spatial-temporal coherence; strokes

摘要: 提出了一种交互式的时空一致性视频抠取方法,能够有效地从视频序列中提取出移动的前景目标.只需在视频体上简单地勾画时空的前景和背景线条,然后构造一个拉普拉斯方程驱动线条在三维的视频体上进行扩散,得到一个粗糙的抠取结果.随后,采用一种新颖的保持时空一致性的抠取技术,利用局部的统计信息和邻域信息,通过少数的迭代即可收敛到全局最优的抠取结果.最后,最优化一个新的全局代价函数在整个三维体上重建前景颜色,如实地保留了抠图结果的时空连贯性.算法的每一步计算都可转化为线性方程组进行求解,因此对于千万像素级的视频数据,也能快速得到高质量的抠取结果.通过对复杂的视频序列进行测试,展示了高质量的抠图结果和算法的高效性.

关键词: 视频抠取;时空一致性;线条矩

图像和视频抠取在电影特技制作、婚纱图像处理等娱乐产业扮演着重要的角色,该技术能够把新的元素无缝地合成到原始场景中,或者赋予导演变换演员位置的能力.视频抠取的关键操作是从背景中分离出移动的前景目标,并且对每一个像素赋予一个估计的前景颜色值 F 以及不透明度值 α . 对于一个给定的图像 I , 该问题等价

^{*} Supported by the National High-Tech Research and Development Plan of China under Grant No.2006AA01Z314 (国家高技术研究发展计划(863))

Received 2008-04-20; Accepted 2008-11-14

于寻找一组最佳的前景颜色 F 、背景颜色 B 以及不透明度值 α ,使得满足如下的形式:

$$I = \alpha F + (1 - \alpha)B \quad (1)$$

对于上述方程,抠图是一个约束不足的逆向问题.早期的“蓝屏”^[1]抠图技术使用单色的背景来简化抠图问题.然而,由于对背景颜色的限定,蓝屏抠图不能处理自然背景的图像.随后,基于三分图(trimap)的前景抠取方法^[2-6]应运而生,得到了令人满意的单幅自然图像的抠取结果.但是,对于一段视频序列逐帧地应用自然图像抠取方法会导致两个问题.首先,对一段较长的视频逐帧地指定三分图非常耗时.尽管光流算法^[7]能够在一定程度上可以加速这一过程,但是当需要处理的视频目标包含快速的移动、剧烈变形的轮廓或者频繁改变的拓扑结构时,双向传播三分图的方法^[8]无法得到良好的结果.另一方面,相邻帧的细小区别将会导致时间域上的不连续性,并产生视觉上的闪烁.

近期,基于硬件的前景抠取方法^[9,10]能够对图像或者视频快速地获得较好的抠取结果.由于这些硬件装置在同一时间只能处理一张图像,因此并没有考虑视频的时空一致性^[11].此外,基于硬件的抠图方法难以处理普通用户拍摄的图像或视频.

我们注意到交互式图像前景抠取的一些新的发展^[12-14],用户通过描绘一些简单的线条粗略地指定前景和背景区域来引导前景抠取.但是,对于视频抠图,逐帧地指定线条仍是一个繁琐和耗时的过程.通过利用帧与帧之间的连贯性,视频剪切技术^[15,16]能够有效地从视频序列中提取出移动的前景目标.然而,由于采用了二元的图切割(graph cut)算法,这些方法只能处理具有光滑前景边界的图像和视频序列.因此,从复杂的自然场景中提取出时空一致的抠图结果,仍是一个充满挑战的难题.

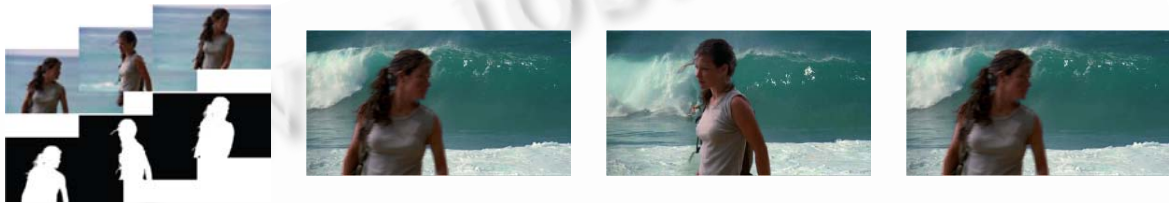


图1 我们的方法能够很方便地从视频中提取出前景目标
左:从输入视频中选取的3帧画面和得到的 α 结果.右:将提取的前景合成到新背景上

在本文,我们提出了一种新的基于线条的交互式视频前景抠取方法,其具有如下的优点:

- 我们提出了一种新的体扩散方式,使得用户指定的时空线条可以快速地在视频体上进行传播.我们可以根据扩散的结果得到一个初始的前景抠取结果,其可以作为下一步抠图的快速预览.
- 我们在整个视频体上执行不透明度值 α 的求解和前景的重建,以保持时空一致的连贯性.
- 我们的方法在体扩散、 α 求解和前景颜色重建的过程中,通过构建二次能量函数,并将其转化成稀疏线性方程组来求快速求解全局最优值.因此,我们的方法能迅速处理千万级的像素数据,如一段 $640 \times 480 \times 100$ 的视频序列.

本文的内容组织如下:首先,我们回顾近年来图像和视频抠图的最新进展.其次,详细地阐述时空一致的交互式视频抠图系统框架.包括基于线条的体扩散、迭代的 α 提取以及时空一致的前景重建3个部分.随后,我们展示视频抠图的实验结果,并对其进行分析和比较.最后,对本文进行总结.

1 相关工作

图像抠图.从背景图像中提取高质量的 α 和前景目标在本质上是一个约束不足的问题.关键在于如何包括尽可能多的先验信息,从而得到近似最优解.

近年来,为了减少用户的交互,BP 抠图^[12]用线条取代三分图,采用迭代的优化框架使线条的先验信息逐步向未知区域扩散.随后,闭合形式抠图^[13]和便捷抠图^[14]将 α 的优化转化为二次代价函数,速度更快,效果更好,进一步发展了基于线条的交互模式.

Wang 等人引入了一种最优的采样选择方案^[18].该方法从已知的前景区域和背景区域内分别采集一组前景样本和背景样本,并对每一采样对进行可靠度分析,只有大于一定可靠度阈值的采样对才能最终被用于抠图能量函数.而实时抠图^[19]则在最优采样选择方案的基础上,提出了一种新的抠图算法.该方法由用户使用笔刷沿着前景目标的边界指定未知区域.通过对边界的分析,笔刷会自动调整粗细.而且实时抠图可以有效地估计前景的颜色,因而使得前景的抠取可以在用户指定笔画的同时进行实时的反馈.

视频抠图.将已有的图像抠图算法扩展到视频序列引起了更多的挑战.这是因为视频抠图对时间连贯性和颜色的误匹配非常敏感.因此,从用户指定的标记和颜色统计中得到的先验信息很可能是错误的.

Zitnick 等人^[20]利用基于分割的立体视觉技术构建了一个视频抠图系统.该系统根据所拍摄的视频序列的深度不连续性来构建三分图.但是,和大多数立体视觉技术一样,它不能处理具有高光表面或是强反射的场景.此外,它要求大量的相机在一个限定的视点范围内采集到高质量的视频数据.

聚焦视频抠图^[9]是第 1 个能全自动地处理动态场景的视频抠图系统.该设备由 3 个同步摄像机所组成,每个摄像机都有不同的聚焦平面.由于前景目标和背景目标存在一定的距离差,使得采集到的 3 段视频流在前景和背景区域有明显的不同,系统计算出前景和背景的大致区域,然后对不确定区域进行优化,计算出三分图,并重建出视频前景.由于深度不连续性的要求,这种方法不能处理前景和背景距离比较接近的场景.

阵列相机视频抠图^[10]使用一个相机阵列同步地采集视频数据.每个相机所采集到的颜色被投影到前景目标的深度平面.随后,统计这些颜色的均值和方差,系统会自动地计算出三分图和前景.和其他基于硬件的抠图系统类似,阵列相机视频抠图需要假定前景和背景之间的颜色距离的方差要足够大.显然,该方法不能很好地处理背景和前景颜色相似的情况.并且如果相机校正的精度不够高,会极大地影响最终的抠图质量.

2 本文的方法

对于各种不同的视频编辑任务^[21],将视频序列视为三维的视频体已经成为一种标准的做法^[22-25].对于视频抠图,我们需要解决两个难点:如何将用户指定的线条快速地传播到整个视频体上;如何保持整个抠图结果的时空一致性.我们的系统框架将上述两个难点转化为能量最小化问题进行统一求解:

$$E = \sum_{p \in \Omega} (E_0(p) + \lambda E_1(p)) \quad (2)$$

Ω 为待求解区域, E_0 和 E_1 分别解决 α 值的准确性(抠图方程的零阶连续性)和抠图结果的时空一致性(抠图方程的一阶连续性). λ 是用户调整的参数,用来平衡上述两项的影响.

用户首先在视频体上指定线条,视频体可以被任意旋转、切割以及切片.一旦结束画线,系统立即对视频体进行全局优化,迅速得到一个初始的 α 值.在初始分割的基础上,通过探索像素的局部统计信息以及空间和时间上的邻域信息,我们在整个视频体上执行三维的时空 α 抠图.最后,根据全局优化函数,我们重建时空一致的前景.在整个过程中,用户可以随时增加线条对提取出的 α 值和重建的前景进行局部精化.

2.1 交互式的体扩散

对于视频抠图,一张好的三分图对于抠图结果的质量是至关重要的.我们系统的目标之一是建立一个简易且高效的界面,使得用户可以便捷地指定前景区域、背景区域以及未知区域.与以往由用户逐帧制作三分图的算法的不同之处在于,我们的系统允许用户在视频体的任意时空曲面上勾画前景和背景线条,如图 2 所示.随后,我们驱动线条所包含的信息在整个视频体上进行扩散,得到一个初始的视频前景抠取结果.

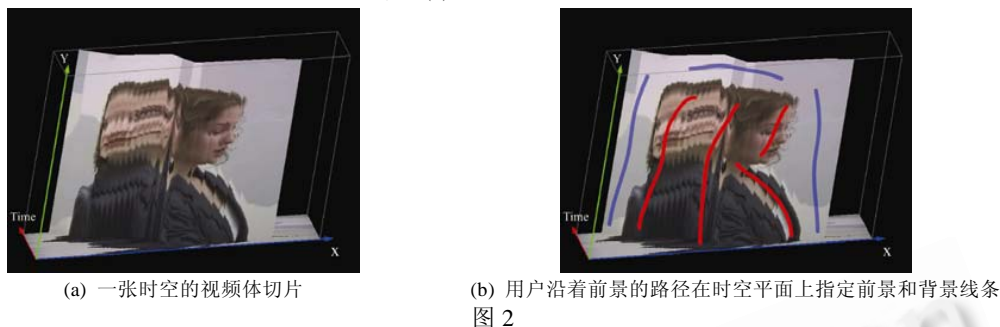
在初始阶段,除了用户勾画线条所指定的那些像素外,其他所有像素的 α 值都是未知的.我们引入体扩散方案是基于以下的两种观察:

(1) 在前景线条和背景线条之间的像素很有可能属于过渡区域,即 α 值在 0 和 1 之间.而剩下的像素则属于前景或背景.

(2) 相邻的像素在颜色空间上很有可能聚为一类.例如,被前景线条所覆盖的像素,和它们相邻并且颜色也非常类似的像素也极有可能属于前景,背景线条也是如此.

根据初始的未知区域 Ω_e , 我们构建六连通的视频体的图结构. 考虑先前提到的对 α 值的两个先验期望以及能量方程(2), 我们定义了如下的体扩散能量方程:

$$E_e = \sum_{p \in \Omega_e} \sum_{q \in N(p)} (\alpha_p - \alpha_q)^2 / \|I_p - I_q\| \quad (3)$$

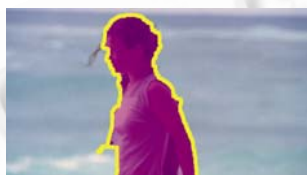


(a) 一张时空的视频体切片

(b) 用户沿着前景的路径在时空平面上指定前景和背景线条



(a) 输入的图像, 线条由用户指定



(b) 由图切割(graph cut)算法进行分割后, 在分割的边界上进行膨胀自动得到的三分图. 所得到的未知区域(白色部分)不能覆盖细长的毛发



(c) 直接应用 Bayesian 抠图算法对(b)进行计算所得到的错误结果



(d) 我们的方法所得到的结果

图 3

在公式(3)中, $N(p)$ 表示像素 p 的邻域. 而 $\|\cdot\|$ 则是计算颜色在 RGB 空间上的距离. 二次方程 $(\alpha_p - \alpha_q)^2$ 则用来衡量相邻像素之间 α 值的差距, 并驱动线条在视频体内扩散. 如果我们简单地令扩散能量为相邻像素之间 α 值的差距的总和, 即 $\sum_{p \in \Omega_e} (\alpha_p - \alpha_q)^2$, 那么未知区域内的所有 α 值会产生过光滑的结果. 所以, 我们采用颜色梯度对能量进行自适应的调整, 即添加系数因子 $1/\|I_p - I_q\|$. 因此, 体扩散能量驱动线条在低对比度区域进行扩散, 而在高对比度区域阻止其扩散行为. 如此, 可以如实地保留前景和背景过渡区域内的细节特征.

我们对公式(3)的两边同时对 α_p 求导:

$$E'_e = \sum_{p \in \Omega_e} \sum_{q \in N(p)} \frac{2\Delta\alpha}{\|I_p - I_q\|} \quad (4)$$

在公式(4)中, $\Delta = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$ 是三维拉普拉斯算子. 上述方程可以构建为线性方程组, 采用双共轭梯度法可对其进行快速求解.

求解能量扩散方程(4)我们可以得到每个像素初始的 α 值, 在此基础上, 整个视频体可分成 3 部分. α 值接近 1 或 0 的像素被视为前景或背景. 剩余的像素归入下一步的计算. 在实验中, 未知区域定义如下:

$$\Omega = \{p \in \Omega_e \mid 0.03 < \alpha_p < 0.97\}.$$

通过图 3 和图 4 的对比,我们展示了体扩散的良好性能.体扩散在我们的系统中扮演了至关重要的角色.



图 4 根据指定的线条以及线条的分布情况所得到的扩散结果,展示了扩散算法的鲁棒性.(a)是和图 5 相同的输入图像.(b)是扩散之后的结果.(c)是由线条扩散所转化而来的三分图.(d)是 α 结果

2.2 迭代的 α 提取

α 的求解过程是一个能量最小化的过程,是一个自适应调整 α 的零阶连续和一阶连续的相互关系的过程.根据方程(2),权重系数 λ 对于同一段视频都是固定的,不同的视频取不同的值(通常取1.0).当前景和背景的对比如比较低的时候, λ 应当取较小的值,反之亦然.

我们假设可以从用户所标记的前景颜色和背景颜色中估计出未知区域的 α ,并以此来定义 E_0 .文献[13]所提出的采样方法在图像抠图上取得了良好的效果,但计算复杂度较高,对于一段数十帧甚至上百帧的视频,计算量非常庞大.因此我们采用文献[14]中所提到的就近采样法.我们的一阶能量项定义如下:

$$E_0(p) = \sum_{i=1}^K \sum_{j=1}^K \|I_p - \alpha_p F_p^i - (1 - \alpha_p) B_p^j\|^2 / \sigma_p^2 \quad (5)$$

颜色的估计 $\alpha_p F_p^i - (1 - \alpha_p) B_p^j$ 是 p 的邻域中的一个最优估计^[18].因此,我们分别选择 K (系统中设为5)个在颜色空间上最接近 I_p 的前背景采样点作为上述颜色估计的采样对. σ_p^2 是颜色估计值和 I_p 距离的方差.

α 的一阶导数可以由 Poisson 抠图^[5]所提出的方法进行近似:

$$\nabla \alpha \approx \frac{\nabla I \cdot (F - B)}{\|F - B\|^2} \quad (6)$$

$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$ 是三维梯度算子.将公式(6)应用到未知区域的所有像素上,我们可以得到视频体上的一个近似的 α 梯度场.为了同步地保持 α 的光滑过渡以及最小化近似的 α 梯度 $\nabla I_p \cdot (F - B)^2 / \|F - B\|^2$ 和所估计的梯度 $\nabla \alpha_p$ 之间的差异,我们提出了如下的一阶项:

$$E_1(p) = \sum_{q \in N(p)} \left(\alpha_p - \alpha_q - \frac{\nabla I \cdot (F - B)}{\|F - B\|^2} \right)^2 \quad (7)$$

因为 F_p 和 B_p 在方程(7)中是未知的,所以我们用离 p 点最近的一对前景和背景采样对 F_p^n 和 B_p^n 代替.我们在公式(2)的两边,对 α_p 进行求导,得到:

$$E'_m = \sum_{p \in \Omega_u} \left(2 \sum_{i=1}^K \sum_{j=1}^K \left((I_p - B_p^j) \cdot (B_p^j - F_p^i) + \|B_p^j - F_p^i\|^2 \alpha_p \right) / \sigma_p^2 + 2\lambda \left(\Delta \alpha_p - \frac{\Delta I_p \cdot (F_p^n - B_p^n)}{\|F_p^n - B_p^n\|^2} \right) \right) \quad (8)$$

方程(8)可以被重写成线性方程组.求解该线性方程组,能够得到时空一致的 α 提取结果.

2.3 时空一致的前景重建

与 α 提取相类似,我们也是在视频体的构架下重建未知像素的前景颜色.在此,方程(2)的第1项 E_0 用来最小化我们所观察到的颜色 I_p 和用前景 F_p ,背景 B_p 以及 α 值 α_p 所重建出来的颜色之间的差异:

$$E_0(p) = \|I_p - F_p - (1 - \alpha)\overline{B}_p\|^2 / \sigma_f^2 \quad (9)$$

α_p 和 \overline{B}_p 分别为上一步计算所得到的 α 值和背景采样点的颜色均值, σ_f^2 是颜色差异的方差,在系统中设为0.5.

视觉上的连贯性^[26]不仅仅对 α 是必须的,而且也是对所需重建的前景颜色的内在要求.我们采用光滑能量项去驱动前景颜色在低对比度区域光滑过渡,而在高对比度区域突出不连续性:

$$E_1(p) = \sum_{q \in N(p)} (F_p - F_q)^2 / (\|I_p - I_q\| + c) \quad (10)$$

这里的 c 是一个常数,用来避免光滑项的奇异性.通常,我们将系数 λ 设为0.1.在公式(2)两边对 F_p 进行求导,我们得到:

$$E'_f = \sum_{p \in \Omega_u} \left(2\alpha \|I_p - \alpha_p F_p - (1 - \alpha_p)\overline{B}_p\| / \sigma_f^2 + 2\lambda \frac{\Delta F_p}{(\|I_p - I_q\| + c)} \right) \quad (11)$$

最小化方程(11)可以生成3个线性方程组,分别是R,G,B这3个颜色通道.

3 实验与结果

文中所有的实验均是在2.8GHz CPU以及2.0GB内存的PC机上进行测试.

在实验中,视频“Kate”的困难之处在于前景目标的快速移动以及经常变化的拓扑结构.视频“Frodo”展示了快速变形的目标轮廓和剧烈抖动的长发.

表1 每段视频的测试数据,时间单位为分钟

视频	尺寸	I	II	III	交互量	关键帧数目
Frodo	580×240×116	1.8	25.0	4.0	45.0	10
Kate	640×350×62	1.4	11.0	2.0	40.0	10
Amira	640×480×79	1.6	10.0	2.0	10.0	8
Elizabeth	504×206×79	3.0	10.0	2.0	10.0	6

表1概括了视频的尺寸、视频抠图系统的运行时间、用户交互量以及指定的键帧的帧数.可以看出,即使是只有少量的用户交互,我们的方法也已展示了高效性和鲁棒性.对于每张关键帧,用户大约需要指定4~5条前景和背景线条.

为了评估我们的算法、Bayesian抠图、便捷抠图以及Robust抠图算法的性能,我们对“Kate”和“Frodo”两段视频序列进行了对比测试.图5展示了视频“Frodo”的对比结果.所测试的两段视频的前景目标均包含快速的移动和剧烈变化的轮廓.逐帧画三分图是一项非常繁琐和耗时的工作,上述两段视频的交互量分别耗时60分钟和120分钟.而我们的方法最多只需在关键帧上画7笔线条.此外,我们的方法计算速度快.根据测试结果,上述优势使得视频抠图的性能比逐帧的图像抠图^[13,14]有了极大的提高(见表2).而逐帧使用便捷图像抠图算法和Robust抠图算法求解视频序列,可以得到好的抠图结果,但帧与帧之间缺乏时空连贯性,抠图结果仍可能会出现视觉上的闪烁.毫无疑问,我们的抠图方法可以在时、空上都较好地保持抠图结果的光滑过渡.

表2 视频“Kate”和“Frodo”的性能测试的对比结果,时间单位为分钟

方法	Bayesian抠图	便捷抠图	Robust抠图	本文的方法
交互方式	三分图	线条	线条	线条
Kate	103.3	93.0	75.0	54.4
Frodo	183.7	125.7	105.0	75.0

用Bayesian视频抠图^[8]对视频“Amira”进行测试,在我们的测试平台上大约耗时150分钟:每帧关键帧的三分图制作需时2分钟,每帧的光流计算、三分图插值以及 α 的提取大约花费90秒.而我们的方法总共只需30分钟即可得到高质量的视频抠图结果(如图7所示).

在图 6 中,我们和 Xue 等人的测地线抠图法^[17],一种当前最快的视频抠图算法之一进行对比.虽然该方法运行效果更高(2 分钟对 25 分钟),并且需要更少的用户输入(2 个关键帧对 6 个关键帧),但是抠图结果显示我们的抠图系统提取出了更高质量的 α .由于 Xue 的方法更加适用于具有硬边界的前景目标的剪切,因而对于多毛的前景目标,其抠图结果会丢失大量的细节.

4 结 论

由于视频目标的内在特性,如大量的数据、复杂的目标运动和视点的变换,使得视频抠图的难度远远大于图像抠图.我们提出的时空一致的视频抠图为用户提供了一个便捷的抠图系统,使其从繁琐的三分图制作过程中解脱出来,从而使视频抠图更加地实用.我们的系统展示了基于线条的简单友好的交互界面,保持了视频抠图的时空连贯性.然而视频前景抠取本身是建立在一定约束条件下的逆向问题,本文的方法虽然在这个问题上取得了一定的进展,但仍存在一些局限.当前景颜色和背景颜色非常接近时,或者颜色剧烈变化时,目前仍很难从输入的视频序列上估计出准确的 α 值.如何改造出更精确的颜色模型,以及如何提供更加方便的局部修改工具,是今后研究的重点.并且我们希望能通过 GPU 加速本文的算法,提高计算效率.

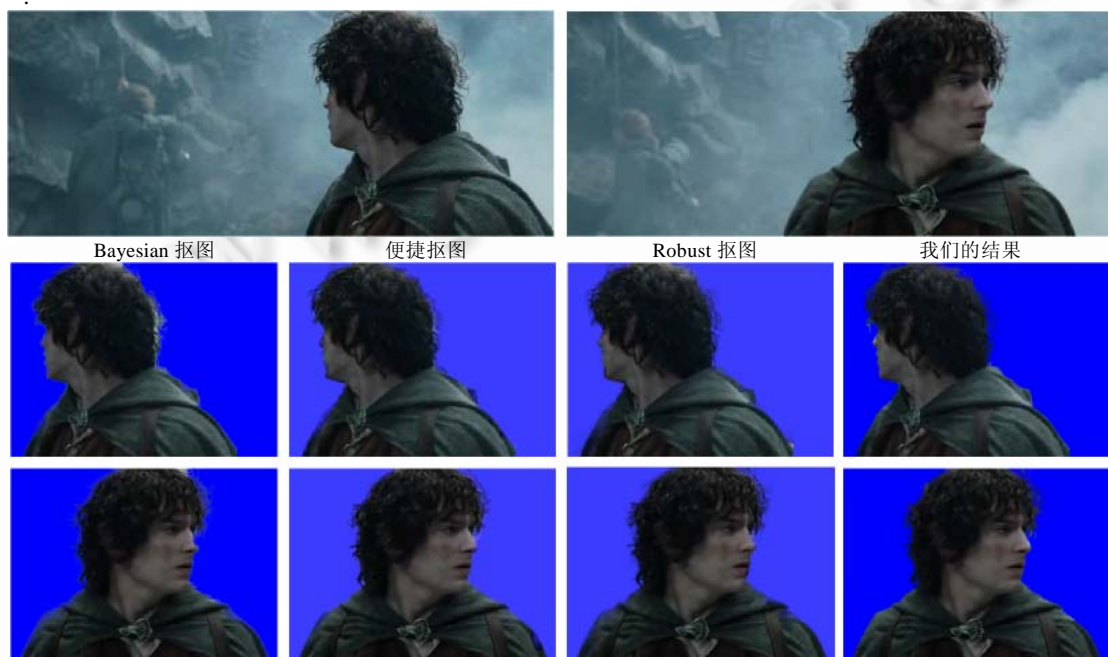


图5 从视频“Frodo”中挑选出的两帧的对比结果. Bayesian 抠图、便捷抠图以及 Robust 抠图产生了视觉缺陷



图6 视频“Elizabeth”结果对比

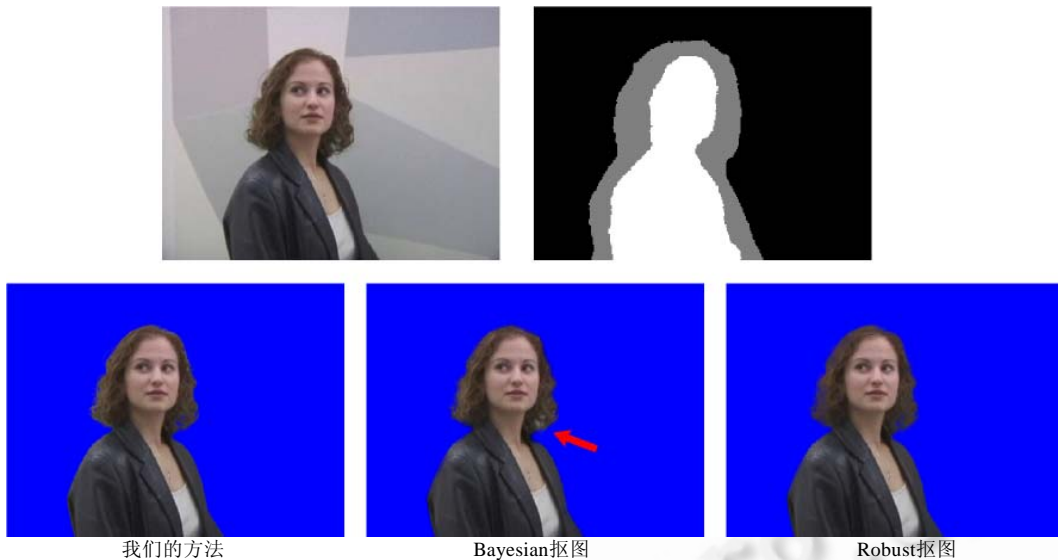


图7 视频“Amira”单帧结果对比.注意到,Bayesian抠图出现了明显的视觉缺陷,如箭头所示

致谢 在此,我们向对本文的工作给予支持和建议的同行,尤其是浙江大学CAD&CG国家重点实验室彭群生教授领导的课题组上的同学和老师表示感谢.

References:

- [1] Smith MA, Blinn JF. Blue screen matting. In: Proc. of the ACM SIGGRAPH 1996. New Orleans, 1996. 259–268.
- [2] Ruzon MA, Tomasi C. Alpha estimation in natural images. In: Proc. of the Computer Vision Pattern Recognition 2000. Hilton Head Island, 2000. 18–25.
- [3] Hillman P, Hannah J, Renshaw D. Alpha channel estimation in high resolution images and image sequences. In: Proc. of the Computer Vision and Pattern Recognition 2001, Vol.1. Hawaii, 2001. 1063–1068.
- [4] Chuang YY, Curless B, Salesin DH, Szeliski R. A Bayesian approach to digital matting. In: Proc. of the Computer Vision and Pattern Recognition 2001, Vol.2. Hawaii, 2001. 264–271.
- [5] Sun J, Jia J, Tang CK, Shum HY. Poisson matting. In: Proc. of the ACM SIGGRAPH 2004. Los Angeles, 2004. 315–321.
- [6] Grady L, Schiwietz T, Aharon S, Westermann R. Random walks for interactive alpha-matting. In: Proc. of the Int'l Conf. on Visualization, Imaging and Image Processing. 2005. 423–429.
- [7] Black MJ, Anandan P. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 1996,63(1):75–104.
- [8] Chuang YY, Agarwala A, Curless B, Salesin DH, Szeliski R. Video matting of complex scenes. In: Proc. of the ACM SIGGRAPH. 2002. 243–248.
- [9] McGuire M, Matusik W, Pfister H, Hughes JF, Durand F. Defocus video matting. *ACM Trans. on Graphics*, 2005,24(3):567–576.
- [10] Joshi N, Matusik W, Avidan S. Natural video matting using camera arrays. *ACM Trans. on Graphics*, 2006,25(3):779–786.
- [11] Apostoloff N, Fitzgibbon AW. Bayesian video matting using learnt image priors. In: Proc. of the Computer Vision and Pattern Recognition, Vol.1. 2004. 407–414.
- [12] Wang J, Cohen MF. An iterative optimization approach for unified image segmentation and matting. In: Proc. of Int'l Conf. on Computer Vision, Vol.2. 2005. 936–943.
- [13] Levin A, Lischinski D, Weiss Y. A closed form solution to natural image matting. In: Proc. of the Computer Vision and Pattern Recognition, Vol.1. 2006. 61–68.
- [14] Guan Y, Chen W, Liang X, Ding ZA, Peng QS. Easy matting—A stroke based approach for continuous image matting. *Computer Graphics Forum*, 2006,25(3):567–576.

- [15] Wang J, Bhat P, Colburn RA, Agrawala M, Cohen MF. Interactive video cutout. ACM Trans. on Graphics, 2005,24(3):585-594.
- [16] Li Y, Sun J, Shum HY. Video object cut and paste. ACM Trans. on Graphics, 2005,24(3):595-600.
- [17] Xue B, Guillermo S. A geodesic framework for fast interactive image and video segmentation and matting. In: Proc. of the Int'l Conf. on Computer Vision. 2007. 584-591.
- [18] Wang J, Cohen MF. Optimized color sampling for robust matting. In: Proc. of the Computer Vision and Pattern Recognition. 2007. 18-25.
- [19] Wang J, Agrawala M, Cohen MF. Soft scissors: An interactive tool for realtime high quality matting. ACM Trans. on Graphics, 2007,26(3):9-16.
- [20] Zitnick CL, Kang SB, Uyttendaele M, Winder S, Szeliski R. High-Quality video view interpolation using a layered representation. ACM Trans. on Graphics, 2004,23(3):600-608.
- [21] Wang HC, Raskar R, Ahuja N. Seamless video editing. In: Proc. of the Int'l Conf. on Pattern Recognition, Vol.3. 2004. 858-861.
- [22] Fels S, Mase K. Interactive video cubism. In: Proc. of the Workshop on New Paradigms for Interactive Visualization and Manipulation. New York, 1999. 78-82.
- [23] Klein AW, Sloan PPJ, Finkelstein A, Cohen MF. Stylized video cubes. In: Proc. of the ACM SIGGRAPH/Eurographics Symp. on Computer Animation. New York, 2002. 15-22.
- [24] Bennett EP, McMillan L. Proscenium: A framework for spatio-temporal video editing. In: Proc. of the ACM Multimedia. 2003. 177-184.
- [25] Wang J, Xu YQ, Shum HY, Cohen MF. Video tooning. In: Proc. of the ACM SIGGRAPH. 2004. 574-583.
- [26] Levin A, Lischinski D, Weiss Y. Colorization using optimization. In: Proc. of the ACM SIGGRAPH. 2004. 689-694



丁子昂(1984-),男,浙江宁波人,硕士生,主要研究领域为图像与视频处理,可视化.



陈为(1976-),男,博士,副教授,主要研究领域为可视化.



夏佳志(1984-),男,硕士,主要研究领域为图像处理,可视化.



彭群生(1947-),男,博士,教授,博士生导师,主要研究领域为真实感图形,数字几何处理,虚拟现实,红外成像仿真,生物计算.



管宇(1981-),男,博士,主要研究领域为图像与视频处理.