

# 一种基于形态学滤波的视频对象空间分割新策略

王煜坚 高建坡 吴镇扬

(东南大学信息科学与工程学院, 南京 210096)

**摘要** 视频对象提取是现代信息科学领域的一个研究热点。为了改进常见的视频对象时空分割方法,提取具有理想边缘的视频对象,本文提出了一种新颖的结合交替顺序重建滤波和自适应阈值判别的空间分水岭分割策略。该策略中结构元素逐渐变大的形态学开闭重建迭代使得分水岭分割所获区域数大大减少,有效地避免了复杂的区域融合。同时,迭代过程使得对象边缘像素的梯度与前景、背景中平坦区域内部像素的梯度更易区分,从而能够应用自适应阈值算法进一步消除由局部梯度极值造成的分割小区域。应用基于该策略的时空分割方法对标准测试序列进行分割实验,实验结果表明,该策略能够带来令人满意的空间分割结果,有助于提高视频对象时空分割方法的主观性能。此外,该策略能够自动确定交替顺序重建滤波器中结构元素的大小以及非线性判别的阈值,算法的通用性和易用性较好。

**关键词** 视频对象提取 分水岭分割 交替顺序重建滤波 自适应阈值

**中图法分类号:** TN911.73 **文献标识码:** A **文章编号:** 1006-8961(2007)06-1018-07

## A Novel Video Object Spatial Segmenting Strategy Based on Morphological Filtering

WANG Yu-jian, GAO Jian-po, WU Zhen-yang

(School of Information Science and Engineering, Southeast University, Nanjing 210096)

**Abstract** Video object extraction has been a most discussed topic in the field of modern information processing. In order to improve the performance of the common temporal-spatial segmenting algorithms and to obtain video objects with ideal boundaries, a novel video object spatial watershed segmenting strategy is proposed in this paper. This strategy combines the merits of alternating sequential filtering by reconstruction and adaptive threshold. The reduction of the number of the segmented regions without complicated area merging is guaranteed by the iteration of morphological filtering of opening and closing by reconstruction with gradually expanding structural elements. Besides, the iteration makes it easier to distinguish the gradient of the object boundary pixels and that of the pixels inside the foreground and background plateaus. This process facilitates the application of the adaptive threshold algorithms to remove the small regions of local gradient extreme. Several comparative temporal-spatial segmenting experiments have been implemented on some standard testing sequences. The experimental results indicate that the proposed strategy has achieved satisfactory spatial segmentation and enhanced the performance of the temporal-spatial segmenting algorithm for the subject and the object. Additionally, appropriate structural element size of the morphological filter and the non-linear threshold values can be automatically generated. This ensures fairly good versatility and usability of the algorithm.

**Keywords** video object extraction, watershed segmentation, alternating sequential filtering by reconstruction, adaptive threshold

## 1 引言

视频对象提取隶属于图像分割的范畴。随着

MPEG-4 和 MPEG-7 标准的研究发展,视频对象分割技术作为新一代视频编码、视频检索、互联网多媒体交互等新兴领域的关键技术,成为当前视频研究领域的热点。

**基金项目:**国家自然科学基金项目(60672094)

**收稿日期:**2006-07-02; **改回日期:**2007-02-12

**第一作者简介:**王煜坚(1980 ~ ),男。2002年于南通大学获学士学位,2005年于东南大学获硕士学位,现为东南大学信号与信息处理专业博士研究生。主要研究方向为视频图像处理、计算机视觉。E-mail: medi@seu.edu.cn

视频对象时空分割算法将时间分割标识出的对象区域和空间分割得到的目标对象在一定准则下融合起来得到最终的视频对象,分割效果较好。其中的空间分割部分利用视频帧内各像素的亮度、色彩以及像素之间的连通关系等信息对帧图像进行分割,能够得到较为精确的目标边缘。目前常见的视频对象空间分割方法有阈值算法、经典分水岭算法、改进分水岭算法等。

经典阈值算法是用经验阈值对图像的亮度、色度或饱和度等特征进行判别。凡是特征数值高于阈值的像素就被判定属于前景对象,反之则被判定属于背景。这种方法简单直观,但是采用的阈值与待分割图像的特点有关,通常需要通过人工试验获得,这就使得算法的通用性和易用性受到局限。此外,当前景对象和背景的特征数值接近甚至相同时,就不能通过阈值来进行判断,经典阈值算法就会失效。

分水岭算法<sup>[1]</sup>是另一种经典的空间分割工具。它能够将图像分割成许多纹理和灰度一致的区域,分割结果能很好地反映图像中的轮廓和边界。但由于经典分水岭算法作用于待分割图像的梯度,噪声和量化误差的影响会导致分割结果中存在许多小区域,这就是经典分水岭算法固有的过分割问题。因此,应用分水岭分割算法主要存在两个缺点:对噪声敏感以及过分割区域的融合过程非常复杂耗时。

对经典分水岭算法的改进主要集中在分割图像的预处理和分割后的区域融合两方面。基于多级小波分解的分水岭分割算法<sup>[2]</sup>就是其中的代表之一。这种算法的思想实质是强调整体,淡化细节。具体来说,就是对待分割图像进行多级金字塔小波分解,在获得的低分辨率图像上应用分水岭算法进行分割,然后将分割结果向高分辨率图像进行区域融合投影。为了得到光滑的目标边缘,融合投影过程要求对每一层金字塔图像计算梯度并应用分水岭分割,计算负荷很大。此外,低分辨率图像分割结果的误差对最终分割结果的影响很大,容易导致最终分割结果边缘和实际前景目标边缘不一致。

因此,解决分水岭算法问题的根本途径是在图像初始分割时尽量抑制噪声与目标内部细密纹理的影响,同时保留重要轮廓,合理降低区域数量,避免区域融合或减小区域融合的难度和复杂度。据此,提出了一种新颖的结合交替顺序重建滤波算法和自适应阈值判别算法的空间分水岭分割策略。

## 2 基于交替顺序重建滤波的分水岭分割算法及其不足

重建滤波器属于形态学滤波器的范畴,它能够在对图像进行滤波的同时很好地保持图像的轮廓信息。交替顺序重建滤波(alternating sequential filtering by reconstruction, ASFBR)<sup>[3]</sup>是一种分级形态学滤波方法,其定义如下:如果 $\gamma_n$ 表示结构元素半径为 $n$ 的形态学开闭重建滤波器,则结构元素半径为 $N$ 的交替顺序重建滤波器可以表示为

$$\Gamma_N(\cdot) = \gamma_N(\gamma_{N-1}(\dots\gamma_2(\gamma_1(\cdot)))) \quad (1)$$

较之于直接用半径为 $N$ 的结构元素进行开闭重建滤波,ASFBR引入的图像畸变更小<sup>[4]</sup>。利用由小到大的结构元素对图像逐次迭代地进行开重建滤波和闭重建滤波,能够将小于结构元素的明亮和黑暗平坦区域逐步合并到周围的区域中,生成灰度特征更均匀的低分辨率表示。这就使得图像大大简化,更加易于分割。

基于ASFBR的分水岭分割算法<sup>[5]</sup>的思想实质也是强调整体,淡化细节,与基于多级小波分解的分水岭分割算法异曲同工。即首先运用交替顺序重建滤波器对图像进行简化,获得更加易于分割的图像低分辨率表示,然后在低分辨图像上应用分水岭分割算法。这样就可以大大减少对图像的过分割,从而避免了计算复杂的区域融合。与基于多级小波分解的分水岭分割算法相比,基于ASFBR的分水岭分割算法的优点在于避免了由低分辨率向高分辨率的区域投影融合过程,大大降低了计算负荷。

但是,基于ASFBR的分水岭分割算法也存在不足。结构元素越大,分割出的区域越少,越有利于避免复杂的区域融合。但另一方面,若结构元素过大,则前景目标中与背景灰度相近的部分会被合并到相邻的背景区域中。这种现象就是所谓的“合并问题”,这种现象会在时空分割结果融合时造成严重的错误。因此,应用基于ASFBR的分水岭分割算法不能为了减少分割区域而一味地增大结构元素尺寸。在实际应用中,结构元素半径通常需要通过试验来人工确定一个合适的折衷值,这就使得算法依赖于待分割图像的分辨率等特征。

所以,如何自动获取合适的结构元素尺寸,避免应用过程中的人工交互,是提高算法的通用性和易用性的关键。

### 3 自适应阈值判别算法及其应用

用较大的结构元素对图像进行交替顺序重建滤波可以大大减少分割区域。但由于图像本身灰度分布的不规则以及结构元素尺寸的影响,滤波后图像的梯度图中仍然存在较多的局部“谷底”和“山峰”,这就导致对滤波后的图像进行分割所得结果中仍然存在大量的小区域。在时空分割结果融合获取视频对象时,如果运动检测模板是多个相邻帧时间分割结果的逻辑和而大于视频对象的实际大小,那么大量的小区域,特别是前景对象实际边缘附近的小区域,就会导致分割出的视频对象具有“背景毛刺”,如图 1 中圈出部分所示。

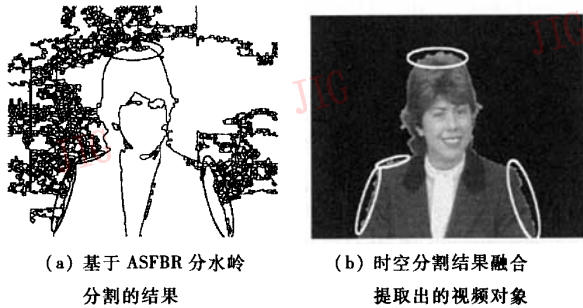


图 1 “背景毛刺”问题示例  
Fig. 1 Illustration of “background burrs” problem

ASFBR 算法可以较好地保持前景目标的边缘而淡化前景与背景内部的细节,这样就使得对象边缘像素的梯度和前景、背景中平坦区域内部像素的梯度之间的差别变大而不产生交叠。这样,就可以在对图像进行交替顺序重建滤波后,应用阈值判别非线性变换来滤除梯度图中的局部“谷底”和“山峰”,进一步去除分割结果中由局部梯度极值造成的小区域,从而解决前景边缘的“背景毛刺”问题。图 2 是对梯度图进行阈值判别变换的示例。

阈值算法的关键在于选取合适的阈值,在一定准则下较好地分出前景和背景。经典阈值算法是对亮度、色度、饱和度等特征进行直方图统计分析,然后通过人工交互确定合适的阈值,这就使得算法难以自动运行。现在提出了多种自适应阈值算法,包括基于最小类内方差的阈值算法<sup>[6]</sup>、基于遗传算法的阈值算法<sup>[7]</sup>等。其中基于最大类间方差的阈值算法<sup>[8]</sup>直观简洁,易于实现。因此本文采用这种算法自动获取阈值,进行非线性判别。

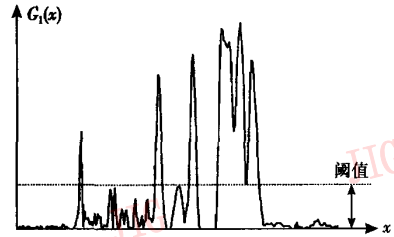


图 2 1 维梯度阈值变换示例  
Fig. 2 Illustration of thresholding 1-D gradient

### 4 结合交替顺序重建滤波和自适应阈值判别的水岭分割新策略

为了自动选取 ASFBR 中结构元素半径  $N$  并解决由局部梯度极值造成大量小区域导致视频对象具有“背景毛刺”的问题,提出一种新颖的结合 ASFBR 算法和自适应阈值算法优点的水岭空间分割策略。算法思想如图 3 所示。

读入待分割视频帧图像后,首先应用梯度算子求取其 2 维梯度。

然后对所得梯度图进行非线性阈值变换。这里采用基于最大类间方差的阈值算法来自动获取判别阈值,其算法表述如下:设 1 帧图像包括  $M$  个像素,

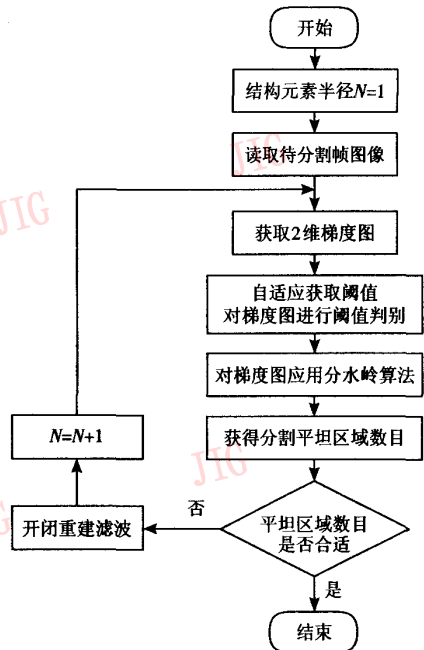


图 3 本文算法框图  
Fig. 3 Diagram of the proposed strategy

其灰度分为  $L$  级  $(0, 1, \dots, L-1)$ 。令  $m_i$  表示灰度为  $i$  的像素总数,  $p_i$  为该灰度级像素在所有像素中的比例。则

$$p_i = \frac{m_i}{M} \quad p_i \geq 0 \quad \sum_{i=0}^{L-1} p_i = 1 \quad (2)$$

取阈值  $t$ , 将所有像素分为两类  $(P_t, P_b)$ ,  $P_b$  包括灰度级  $[0, \dots, t]$  的像素,  $P_t$  包括灰度级  $[t+1, \dots, L-1]$  的像素。则类  $P_b$  和  $P_t$  的灰度均值分别为

$$\mu_b(t) = \frac{\sum_{i=0}^t ip_i}{\omega_b(t)} \quad \text{其中 } \omega_b(t) = \sum_{i=0}^t p_i \quad (3)$$

$$\mu_t(t) = \frac{\sum_{i=t+1}^{L-1} ip_i}{\omega_t(t)} \quad \text{其中 } \omega_t(t) = 1 - \omega_b(t) \quad (4)$$

定义类间方差  $\sigma_R^2(t) = \omega_b(t)\omega_t(t)[\mu_b(t) - \mu_t(t)]^2$ , 其中  $\mu_b(t)$  和  $\mu_t(t)$  分别代表背景和前景目标的中心灰度。要获得最佳阈值, 就要使  $\sigma_R^2(t)$  最大, 即

$$t^* = \arg \max_{0 \leq t \leq L-1} (\sigma_R^2(t)) \quad (5)$$

在具体实现时, 最佳阈值可以通过下式获得

$$t^* = \arg \max_{0 \leq t \leq L-1} (\omega_b(t)\omega_t(t) |\mu_b(t) - \mu_t(t)|) \quad (6)$$

得到最佳阈值  $t^*$  后, 对图像进行判别。如果像素灰度值高于  $t^*$ , 则将该像素判定属于前景或重新出现的背景; 反之, 将该像素判定属于背景。

用上述基于最大类间方差的阈值算法对 2 维梯度图进行处理, 就可以较好地滤除梯度图中的局部极值。

将经过阈值处理的梯度图作为分水岭算法的输

入, 待分割结束后, 统计分割出的区域数目, 然后对所得结果进行判断。图 3 中循环中断条件“平坦区域数目是否合适”是指在不出现在前景背景“合并问题”的情况下使得分割出的平坦区域尽可能少。通过实验分析, 对于自然视频图像, 当所获平坦区域数目为直接对原始图像应用分水岭算法所获平坦区域数目的 5% 左右 (对于静止背景, 前景背景对比较大的视频图像, 甚至更低) 时, 就能获得令人满意的分割结果。当循环中断条件满足时, 就将分割结果输出, 算法结束。否则, 就对帧图像进行一次开重建和闭重建滤波, 然后将结果送回求取梯度的步骤, 同时结构元素半径加 1 进入下一次迭代。

选取 MPEG-4 标准测试视频序列 Claire (CIF) 的第 152 帧进行实验。实验平台配置为 AMD Athlon64 3000 + (2.0GHz) CPU, 768M RAM, MATLAB 7.02。

首先用人工确定大小的结构元素对帧图像进行基于 ASFBR 的分水岭分割, 结果如图 4 所示。

统计分割区域数目, 如表 1 所示。

表 1 基于 ASFBR 的分水岭分割算法所得分割区域数目

Tab.1 Number of the watershed segmented regions based on ASFBR

	分割区域数目
原始图像	8 293
$N=20$	2 903
$N=40$	2 163
$N=50$	1 598

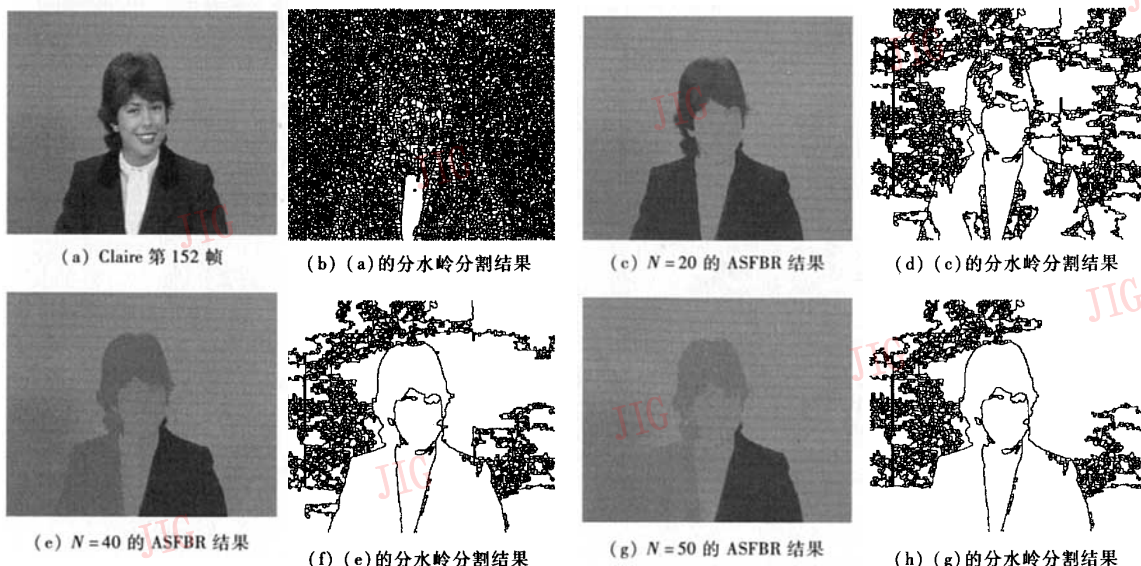


图 4 用人工确定大小的结构元素对 Claire 第 152 帧进行基于 ASFBR 的分水岭分割结果

Fig. 4 Watershed segmentation of Claire's 152nd frame after ASFBR with artificially selected structural elements

显然,  $N$  值越大, 所得分割区域越少。这就能够避免区域融合或减小区域融合的难度和复杂度。

同时, 对滤波后图像的梯度进行考察, 可以发现, 随着迭代中结构元素越来越大, 前景对象边缘像

素的梯度和前景、背景中平坦区域内部像素的梯度越来越易区分, 如图 5 所示。这样就可以应用自适应阈值算法进一步去除由局部梯度极值造成的分割小区域, 准确地提取对象边缘。

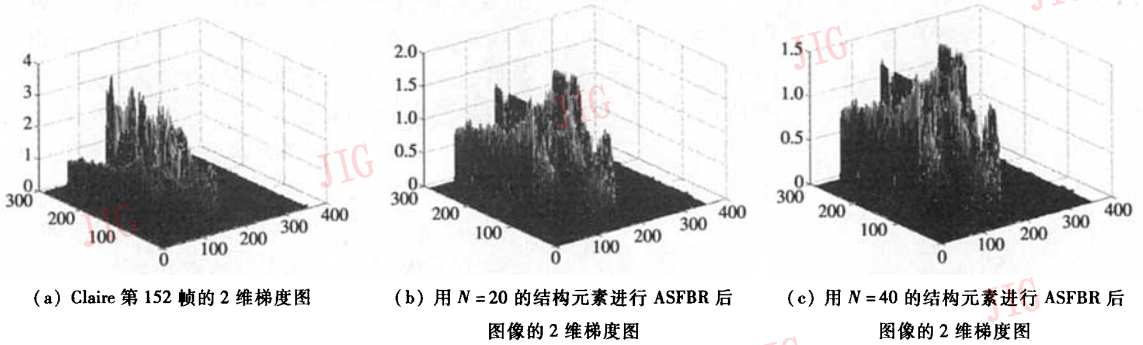


图 5 ASFBR 对图像梯度的作用示例

Fig. 5 Illustration of the effects on the gradient meshes by ASFBR

在图 4(h) 中, Claire 身体的左下部分和背景发生了“合并问题”, 这是因为结构元素半径  $N = 50$  过大了。因此, 为了防止前景对象出现误判的情况,  $N$  必须有一个上限。

应用本文的分割策略对 Claire 第 152 帧进行处理。在每一次迭代过程中, 求取经过 ASFBR 后的图像的 2 维梯度, 然后运用基于最大类间方差的阈值算法去除局部梯度极值, 接着进行分水岭分割操作, 之后统计分割区域的数目并进行判断, 如果区域数目满足循环中断条件, 则将分割结果输出; 否则, 算法进入下一次迭代。这样就能自动确定合适的结构元素大小, 并最终获得理想的空间分割结果。

图 6 为迭代过程中  $N = 20$  时的中间结果。

最终获得的空间分割结果如图 7 所示。

显然, 本文提出的分割策略获得了与采用最佳人工交互的算法相似的空间分割结果, 大大减少了

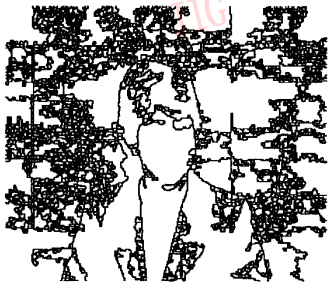
分割所得区域的数目, 有效减小了区域融合负担。

采用合适的时空分割融合方法将上述空间分割结果与适当的时间分割结果结合起来, 就能最终提取出视频对象。图 8 是基于不同空间分割算法的时空分割方法的主观性能对比示例。

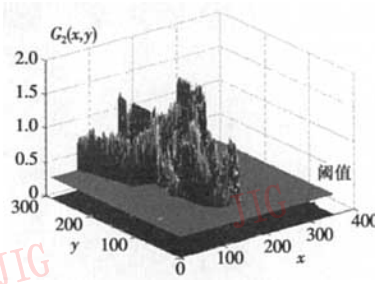
可见, 本文提出的分割策略进一步有效地减少了局部梯度极值造成的小区域, 从而消除了视频对象的“背景毛刺”。此外, 本算法中 ASFBR 的结构元素大小和非线性判断阈值都是自动获得, 无需人工交互, 算法的通用性和易用性较好。

下面对算法的客观性能进行考察。空间准确率<sup>[10]</sup>是常用的客观评价标准之一, 定义如下:

$$SA(t) = d(A_i^{ent}, A_i^{ref}) = 1 - \frac{\sum_{(x,y)} A_i^{ent}(x,y) \oplus A_i^{ref}(x,y)}{\sum_{(x,y)} A_i^{ref}(x,y)} \quad (7)$$



(a) 用  $N = 20$  的结构元素进行 ASFBR 后分水岭分割的结果



(b) 对 2 维梯度图进行自适应阈值变换示意



(c) 对 2 维梯度图阈值判别后的分水岭分割结果

图 6 ASFBR 迭代过程中自适应阈值判别结果示例

Fig. 6 Adaptive thresholding in one circle of the iteration

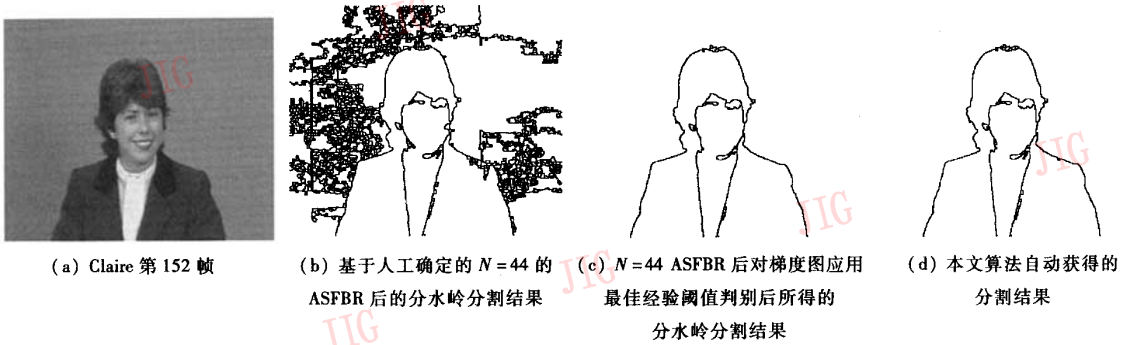


图 7 实验结果与比较

Fig.7 Experiment results and the comparison

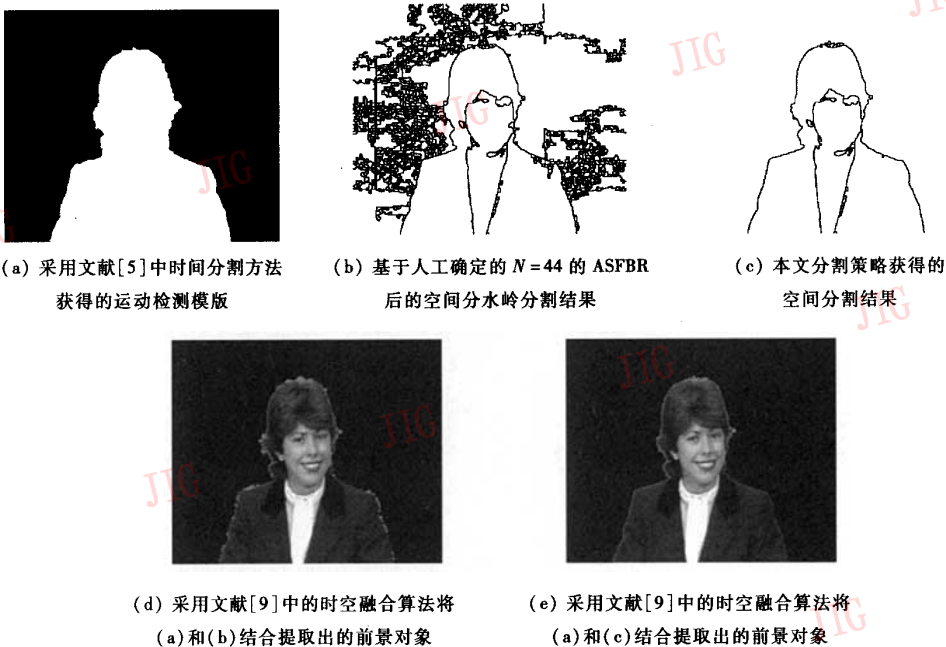


图 8 Claire 序列第 152 帧对比实验结果

Fig.8 Comparative experiment results of Claire's 152nd frame

其中,  $A_t^{act}$  和  $A_t^{ref}$  分别对应第  $t$  帧分割模板和参考模板,  $\oplus$  表示“异或”操作。

以人工分割的图像作为标准参考分割图像(称为 Ground-Truth)。用获得图 8(e)的方法对 Claire 序列第 2 ~ 200 帧进行分割的平均空间准确率为 98.9%, MATLAB 环境下提取一帧对象所需的平均时间为 9.652s。而用获得图 8(d)的方法进行分割的平均空间准确率为 96.8%, 同时算法的有效运行依赖于人工交互。

选择背景复杂度更高的 Akiyo(QCIF)序列进行实验。图 9 是对 Akiyo 序列第 3 帧的对比实验结果。

用获得图 9(e)的方法对 Akiyo 序列第 2 ~ 250 帧进行分割的平均空间准确率为 97.8%, MATLAB 环境下提取一帧对象所需的平均时间为 4.237s。而用获得图 9(d)的方法进行分割的平均空间准确率为 92.1%, 同时算法的有效运行依赖于人工交互。

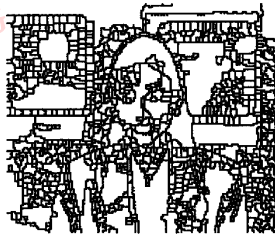
通过上述实验结果可以明显看出,本文提出的空间分割策略有助于时空分割算法获得更好的主客观性能。

### 5 结 论

提出了一种视频对象分水岭空间分割新策略。



(a) 采用文献[5]中时间分割方法获得的运动检测模版



(b) 基于人工确定的  $N=7$  的 ASFBR 后的空间分水岭分割结果



(c) 本文分割策略获得的空间分割结果



(d) 采用文献[9]中的时空融合算法将 (a) 和 (b) 结合提取出的前景对象



(e) 采用文献[9]中的时空融合算法将 (a) 和 (c) 结合提取出的前景对象

图 9 Akiyo 序列第 3 帧对比实验结果

Fig. 9 Comparative experiment results of Akiyo's 3rd frame

本算法结合了交替顺序重建滤波算法和自适应阈值判别算法的优点。在结构元素逐渐变大的迭代过程中,大大减少了所获分割区域的数目,有效避免了复杂的区域融合,同时使得对象边缘像素的梯度与平坦区域内部像素的梯度更易区分,利用基于最大类间方差的阈值算法自动获得的判别阈值,通过非线性变换进一步消除局部梯度极值造成的分割小区域,从而获得令人满意的视频对象空间分割结果。将应用本文分割策略获得的空间分割结果与合适的时间分割结果相结合,就可以有效提取出具有精确边缘的视频对象。此外,本算法中 ASFBR 的结构元素大小和非线性判断阈值都是自动获得,无需人工交互,算法的通用性和易用性较好。

### 参考文献 (References)

- 1 Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, 13(6): 583 ~ 598.
- 2 Yang Gao-bo, Zhang Zhao-yang, Yu Sheng-fa. Automatic video object extraction algorithm based on wavelet and watershed [J]. *Journal on Communications*, 2005, 26(3): 7 ~ 13, 20. [杨高波, 张兆杨, 余圣发. 一种基于小波分解和分水岭变换的视频对象自动分割算法[J]. *通信学报*, 2005, 26(3): 7 ~ 13, 20.]
- 3 Liu Jian-qiu, Ruan Qiu-qi. The research and application of

morphological filter by reconstruction [J]. *Journal on Communications*, 2002, 23(1): 116 ~ 121. [刘剑秋, 阮秋琦. 形态学重建滤波器的研究与应用[J]. *通信学报*, 2002, 23(1): 116 ~ 121.]

- 4 Morales A, Acharya R, Sung-Jea K. Morphological pyramids with alternating sequential filters [J]. *IEEE Transactions on Image Processing*, 1995, 4(7): 965 ~ 977.
- 5 Munchurl K, Jae G C, Daehee K, et al. A VOP generation tool: automatic segmentation of moving objects in mage sequences based on spatio-temporal information [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, 9(8): 1216 ~ 1226.
- 6 Brink A D. Thresholding of digital images using two - dimensional entropies [J]. *Pattern Recognition*, 1992, 28(8): 803 ~ 808.
- 7 Zheng Hong, Pan Li. The automatic selection of image threshold on the basis of genetic algorithms [J]. *Journal of Image and Graphics*, 1999, 4(4): 327 ~ 330. [郑宏, 潘励. 基于遗传算法的图像阈值的自动选取[J]. *中国图象图形学报*, 1999, 4(4): 327 ~ 330.]
- 8 Ostu N. A threshold selection method from graylevel histogram [J]. *IEEE Transaction on Systems, Man and Cybernetics*, 1978, SMC9(1): 62 ~ 66.
- 9 Pan J H, Li S P, Zhang Y Q. Automatic extraction of moving objects using multiple features and multiple frames [A]. In: *Proceedings of IEEE International Symposium on Circuits and Systems [C]*, Geneva, Switzerland, 2000: 36 ~ 39.
- 10 Yang G B, Zhang Z Y. Video object segmentation for head-shoulder sequences in the cellular neural networks architecture [J]. *Real-time Imaging*, 2003, 9(3): 171 ~ 178.