

# 基于层次光流的半自动时空视频分割技术

赵明 陈纯 邬正平

(浙江大学计算机科学与工程系 杭州 310027)

(浙江大学虹软多媒体研究中心 杭州 310027)

**摘要** 在新一代 MPEG-4 视频编码标准中, 为了支持面向对象编码和实现基于内容的应用, 视频的半自动分割成为关键技术之一。本文提出一种基于层次光流的半自动时空视频分割算法。算法由空域分割和时域分割组成。在空域分割中, 本文提出的基于点的图形用户界面(PBGUI)在用户的协助下, 能够精确地定义需要分割的视频对象(VO)。时域分割根据空域分割的结果采用层次光流算法对视频对象进行边界和整体跟踪。实验结果表明, 利用本文提出的算法, 能够较精确地分割出视频对象。

**关键词** MPEG-4, 光流, 视频分割, 面向对象编码

中图法分类号: TP391.4

## hierarchy optical flow based semi-automatic

## spatial-temporal video segmentation

ZHAO Ming CHEN Chun WU Zheng-ping

(Department of Computer Science and Engineering, Zhejiang University, hangzhou, 310027)

**Abstract** In the new MPEG-4 video coding standard, the semi-automatic video segmentation plays a key role in supporting object-oriented coding and enabling content-based functionalities. A novel hierarchy optical flow based semi-automatic spatial-temporal video segmentation method is presented in this paper. The proposed algorithm comprises of spatial and temporal segmentation modules. In the spatial segmentation stage, the user can input points around the video object (VO) with the proposed point-based graphic user interface (PBGUI), then active contour model and tracking bug algorithm are used to precisely define the video object of interest to be segmented. With the result of spatial segmentation, the temporal segmentation involves non-rigid object boundary tracking and rigid object whole entity tracking by hierarchy optical flow algorithm based on the algorithm proposed by Lucas and Kanade. And the tracking points selection algorithm is proposed to greatly improve the tracking performance in the rigid object whole entity tracking. The experimental results show that the proposed algorithm can precisely segment video objects from video clips and can be applied to object-oriented coding, content-based functionality and multimedia database indexing.

**Keywords** MPEG-4, optical flow, video segmentation, object-oriented coding

## 0 引言

随着视频编码技术的发展, 基于对象的编码方案日益成为研究热点。在新一代视频编码标准 MPEG-4 中, 视频序列中的图像被表示为视频对象<sup>[1]</sup>的集合, 不仅大大提高编码效率, 而且使基于内容的交互成为可能。然而, MPEG-4 却假设在编码之前, 视频序列已经被表示为视频对象。所以, 视频对象的获得(即视频分割)对于 MPEG-4 以及其相关特性的实现至关重要。

视频分割技术通常分为自动分割技术和半自动分割技术<sup>[2]</sup>。自动分割技术<sup>[3][4]</sup>通过定义基于颜色、纹理、边缘、运动等低级视觉特征的一致性标准进行分割, 但是研究结果表明自动分割技术要么面向特定应用, 要么不能达到令人满意的通用效果。所以半自动分割技术越来越受到重视。半自动分割技术<sup>[5][6]</sup>通常是用户通过图形用户界面(GUI)对视频图像进行初始分割, 给出视频对象初始的精确轮廓, 然后对后继帧采取自动分割方法进行分割。

文献<sup>[5]</sup>提出一种半自动分割算法, 此算法采用多值分水岭算法进行空域分割, 再利用空域分割结果进行跟踪。但是此算法对非刚体跟踪效果不好, 难以跟踪大运动的非刚体运动, 同时没有对刚体分割进行特别处理。文献<sup>[6]</sup>提出的半自动分割算法分为帧内分割和帧间分割。帧内分割先让用户在视频

对象的边界周围给出一个覆盖边界的带，再采用分水岭算法得到视频对象。然后，帧间分割自动地分割出视频对象。但是，文献<sup>[5]</sup>的空域分割和文献<sup>[6]</sup>帧内分割结果严重依赖于图像分割结果，不便于用户修改，交互性能差。文献<sup>[6]</sup>的帧间分割虽然分为刚体和非刚体分割，但是对非刚体的边界跟踪采用块匹配算法，准确性差；对刚体的区域跟踪又没有考虑到刚体内部点的易跟踪性，所以跟踪性能较差。

本文提出一种基于层次光流的半自动时空视频分割算法，可以克服上述缺点。算法由空域分割和时域分割组成。空域分割采用本文提出的基于点的图形用户界面(PBGUI)，可以很好地实现用户交互，精确定义视频对象(VO)。时域分割对非刚体的边界跟踪和刚体的整体跟踪均采用了基于层次光流的算法，同时提出跟踪点选择算法，极大地提高了分割性能。本文算法与其他视频分割算法相比有以下特点：(1) 独特的基于点的图形用户界面 (PBGUI)；(2) 边界跟踪和整体跟踪均采用层次光流算法。(3) 提出跟踪点选择算法提高了分割效果。

## 1 空域分割

空域分割的目的是在用户协助下得到视频对象的精确轮廓，它必须完成两个功能：第一，便于用户交互的图形用户界面(GUI)。第二，通过用户交互，得到视频对象的精确轮廓。本文提出一种基于点的图形用户界面 (PBGUI)，能够很好地实现上述功能。

**1.1 便于用户交互的图形用户界面 - PBGUI。**一般地说，用户最易于输入大致的点，其次是准确的点，然后是大致的线，最后是准确的线。文献<sup>[6]</sup>采用基于带的方法，实际要求用户输入大致的线，其缺点在于：(1) 要求用户输入较多数据；(2) 用户较难输入大致的线 - “带”。而本文提出的基于点的图形用户界面 (PBGUI)，只要求用户输入大致的点。在 PBGUI 中，用户只需在视频对象边界附近指定一些点，然后系统就可以求出视频对象的轮廓，如果轮廓不够精确，用户只需在不够精确的地方调整已输入的点或是增加一些点，系统就可以给出更精确的轮廓。

**1.2 利用用户交互，得到视频对象的精确轮廓。**文献<sup>[6]</sup>中，用户输入覆盖边界的带，系统通过图像分割和区域分类得到视频对象的轮廓。此方法有两个缺点：(1) 得到的轮廓不便于修改。高效的修改应该只修改不正确之处，同时保留正确的地方，但是这两点此方法都不能做到。因为此方法不能局部地修改用户的输入带，所以，在修改时，用户必须重新输入整个边界带，但是这样却使得原来已经正确的地方没法保留下来。(2) 难于得到精确的边界。因为其边界的准确性完全依赖于图像分割的效果，一旦图像分割效果不理想，就无法修改到正确的边界上。本文提出的 PBGUI 可以很好地解决上述问题。

本文基于 PBGUI 的算法如下：(1) 利用主动轮廓线模型<sup>[7]</sup> (ACTIVE CONTOUR MODEL OR SNAKE) 对用户通过 PBGUI 输入的点进行调整。由于用户输入点在边界附近，所以主动轮廓线模型可以把这些点调整到边界上。称调整后的点为控制点。(2) 用改进的边界跟踪算法，从控制点得到视频对象的边界线。如果起始点选择正确，则整个边界的准确性有较高保证<sup>[8]</sup>。本文采用改进的具有很强抗噪能力的跟踪虫算法<sup>[9]</sup>，对跟踪虫算法进行了如下改进：第一，算法从起始点和终止点同时跟踪，解决了一般跟踪虫算法对起始点过度依赖的问题。第二，限定边界不能急速地变化方向，保证边界光滑。(3) 用区域填充算法，从边界线得到视频对象。本文采用的算法是从图像边界向内填充，不能填充的区域就是视频对象的轮廓。

由上面的分析可见，本文提出的PBGUI有如下的优点：(1) 界面友好，易于输入，输入数据量少。(2) 交互性好，易于修改。(3) 结果精确，能够得到精确的视频对象。所以，本文的算法能够很好地完成空域分割的目的。

## 2 时域分割

时域分割利用空域分割结果，自动分割出视频对象。视频对象分为刚体和非刚体。本文对刚体采用整体跟踪法进行分割，对非刚体采用边界跟踪法进行分割，二者都是采用 Lucas-Kanade 光流模型，不仅理论基础统一，而且分割效率高。特别地，在对刚体的整体分割中，本文提出跟踪点选择算法，

极大地提高了分割效果。

## 2.1 基于边界跟踪的非刚体分割

对非刚体的分割，文献<sup>[6]</sup>采用了块匹配方法进行边界跟踪。但是文献<sup>[10]</sup>指出，块匹配方法不擅长于小运动量的运动估算，因此文献<sup>[6]</sup>通过边界精化修正匹配结果。然而边界精化本身的准确性又难以保证，尤其是在有其他边界干扰的情况下修正效果更差，所以其总体效果并不理想。

文献<sup>[10]</sup>同时指出，在小运动量情况下，Fleet 和 Jepson 的方法总体效果最好，Lucas - Kanade 方法次之，但效率更高。考虑到分割算法必须考虑时间因素，故本文采用 Lucas - Kanade 方法进行边界点跟踪。同时，本文利用高斯层次模型，使得此方法能够适用于大运动量情况。

### 2.1.1 Lucas - Kanade 光流跟踪模型

Lucas - Kanade<sup>[11]</sup>方法假设运动矢量在特定的象素块 B 上保持不变，即  $V(X,t)=[v_x(t) \ v_y(t)]^T=[a \ d]^T \quad X \in B \quad (1)$ ，其中  $X=(x,y)$ 。如果象素块 B 包含足够的灰度等级变化，就能

够唯一估算出运动矢量。虽然这种模型不能处理旋转运动，但是对于局部小块，其运动可以认为是平移运动，同时视频对象边界上的灰度变化相当丰富，所以采用 Lucas - Kanade 方法进行边界点跟踪相当合适。Lucas 和 Kanade<sup>[11]</sup>定义在象素块 B 上的光流方程误差为  $E = \sum_{X \in B} \left( \frac{\partial I(X,t)}{\partial x} a + \frac{\partial I(X,t)}{\partial y} d + \frac{\partial I(X,t)}{\partial t} \right)^2$

(2)，其中， $I(X,t)$  表示图像的亮度函数。为使误差最小，分别对 a、d 计算误差 E 的偏导数，令它们都等于零，得到  $\sum_{X \in B} \left( \frac{\partial I(X,t)}{\partial x} a + \frac{\partial I(X,t)}{\partial y} d + \frac{\partial I(X,t)}{\partial t} \right) \frac{\partial I(X,t)}{\partial x} = 0 \quad (3)$ ，

$\sum_{X \in B} \left( \frac{\partial I(X,t)}{\partial x} a + \frac{\partial I(X,t)}{\partial y} d + \frac{\partial I(X,t)}{\partial t} \right) \frac{\partial I(X,t)}{\partial y} = 0 \quad (4)$ ，联立求解此方程组得

$$\begin{pmatrix} a \\ d \end{pmatrix} = \begin{bmatrix} \sum_{X \in B} \frac{\partial I(X,t)}{\partial x} \frac{\partial I(X,t)}{\partial x} & \sum_{X \in B} \frac{\partial I(X,t)}{\partial x} \frac{\partial I(X,t)}{\partial y} \\ \sum_{X \in B} \frac{\partial I(X,t)}{\partial x} \frac{\partial I(X,t)}{\partial y} & \sum_{X \in B} \frac{\partial I(X,t)}{\partial y} \frac{\partial I(X,t)}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} -\sum_{X \in B} \frac{\partial I(X,t)}{\partial x} \frac{\partial I(X,t)}{\partial t} \\ -\sum_{X \in B} \frac{\partial I(X,t)}{\partial y} \frac{\partial I(X,t)}{\partial t} \end{bmatrix} \quad (5)$$

其中， $\frac{\partial I(X,t)}{\partial x}$ 、 $\frac{\partial I(X,t)}{\partial y}$ 、 $\frac{\partial I(X,t)}{\partial t}$  表示图像中的点 X 在 x 轴、y 轴以及时间轴 t 上的梯度。同时，(5) 式中运动量的计算有一个限制条件，即矩阵

$$\begin{bmatrix} \sum_{X \in B} \frac{\partial I(X,t)}{\partial x} \frac{\partial I(X,t)}{\partial x} & \sum_{X \in B} \frac{\partial I(X,t)}{\partial x} \frac{\partial I(X,t)}{\partial y} \\ \sum_{X \in B} \frac{\partial I(X,t)}{\partial x} \frac{\partial I(X,t)}{\partial y} & \sum_{X \in B} \frac{\partial I(X,t)}{\partial y} \frac{\partial I(X,t)}{\partial y} \end{bmatrix} \quad (6) \text{ 可逆。}$$

### 2.1.2 非刚体分割算法

基于 Lucas - Kanade 光流跟踪模型的非刚体分割算法以当前帧  $F_t$  中视频对象的掩膜 ( $Mask_t$ ) 为输入，以下一帧  $F_{t+1}$  中视频对象掩膜 ( $Mask_{t+1}$ ) 为输出，算法如下：

- (1) 对当前帧  $F_t$  中视频对象掩膜 ( $Mask_t$ ) 边界上的每一点用 (5) 式求出其运动向量。对于无法求出运动向量的点，其运动向量用相邻两点运动向量的插值给出。
- (2) 用当前帧  $F_t$  的边界点和对应的运动向量求出下一帧  $F_{t+1}$  中视频对象的边界点，并将边界点连成封闭边界线。
- (3) 由区域填充算法 (见 2.2) 求出下一帧  $F_{t+1}$  中的视频对象掩膜 ( $Mask_{t+1}$ )。

同时，本文采用高斯层次模型使算法适用于较大运动量，并且增强抗噪了能力。实现层次模型时，在每一层中都运行上面的算法，所不同的是，此层边界点的初始位置是上层边界点的位置与运动向量之和的两倍，最低层结果就是  $Mask_{k+1}$ 。

## 2.2 基于整体跟踪的刚体分割

对于刚体，可以用整体运动表示边界运动，所以采用整体运动模型进行跟踪，不仅能够提高跟踪效果，而且算法效率跟高。本文采用基于 Lucas - Kanade 方法的两参数整体平移模型和六参数整体仿射模型进行整体跟踪。

### 2.2.1 整体平移模型

事实上，(1) 式所表示的就是平移模型。在对非刚体进行分割时，象素块 B 选在边界点上，故运动向量表示边界块的平移运动。如果将象素块 B 修改为视频对象所代表的区域，则计算出的运动向量便表示视频对象的整体平移运动。用  $T$  表示修改后的矩阵 (6)，用  $M_T$  和  $B_T$  分别表示修改后的 (5) 式最左和最右的矩阵，则整体平移模型为： $M_T = [m_t^a, m_t^d]^T = T^{-1}B_T$  (7)

### 2.2.2 整体仿射模型

整体仿射模型采用六参数仿射变换，它假设运动矢量在特定的象素块 B 上满足相同的仿射变换模型，这样 (1) 式变为  $V(X,t)=[v_x(t) \ v_y(t)]^T=[a+bx+cy \ d+ex+fy]^T \ X \in B$  (8)，其中  $X=(x,y)$ 。如整体平移模型，在整体仿射模型中的 B 表示视频对象所代表的区域。同样 (2) 式变为

$$E = \sum_{X \in B} \left( \frac{\partial I(X,t)}{\partial x} (a+bx+cy) + \frac{\partial I(X,t)}{\partial y} (d+ex+fy) + \frac{\partial I(X,t)}{\partial t} \right)^2 \quad (9)$$

f 计算误差 E 的偏导数，令它们等于零，得到一个方程组，联立求解此方程组得

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} \sum_{X \in B} I_x I_x & \sum_{X \in B} x I_x I_x & \sum_{X \in B} y I_x I_x & \sum_{X \in B} I_x I_y & \sum_{X \in B} x I_x I_y & \sum_{X \in B} y I_x I_y \\ \sum_{X \in B} x I_x I_x & \sum_{X \in B} x^2 I_x I_x & \sum_{X \in B} xy I_x I_x & \sum_{X \in B} x I_x I_y & \sum_{X \in B} x^2 I_x I_y & \sum_{X \in B} xy I_x I_y \\ \sum_{X \in B} y I_x I_x & \sum_{X \in B} xy I_x I_x & \sum_{X \in B} y^2 I_x I_x & \sum_{X \in B} y I_x I_y & \sum_{X \in B} xy I_x I_y & \sum_{X \in B} y^2 I_x I_y \\ \sum_{X \in B} I_x I_y & \sum_{X \in B} x I_x I_y & \sum_{X \in B} y I_x I_y & \sum_{X \in B} I_y I_y & \sum_{X \in B} x I_y I_y & \sum_{X \in B} y I_y I_y \\ \sum_{X \in B} x I_x I_y & \sum_{X \in B} x^2 I_x I_y & \sum_{X \in B} xy I_x I_y & \sum_{X \in B} x I_y I_y & \sum_{X \in B} x^2 I_y I_y & \sum_{X \in B} xy I_y I_y \\ \sum_{X \in B} y I_x I_y & \sum_{X \in B} xy I_x I_y & \sum_{X \in B} y^2 I_x I_y & \sum_{X \in B} y I_y I_y & \sum_{X \in B} xy I_y I_y & \sum_{X \in B} y^2 I_y I_y \end{bmatrix}^{-1} \begin{bmatrix} - \sum_{X \in B} I_x I_t \\ - \sum_{X \in B} x I_x I_t \\ - \sum_{X \in B} y I_x I_t \\ - \sum_{X \in B} I_y I_t \\ - \sum_{X \in B} x I_y I_t \\ - \sum_{X \in B} y I_y I_t \end{bmatrix} \quad (10)$$

此式表示视频对象的整体运动模型，分别用  $M_A$ 、 $A^{-1}$ 、 $B_A$  表示 (10) 式中从左到右的三个矩阵，则整体运动模型可以表示为： $M_A = [m_a^a, m_a^b, m_a^c, m_a^d, m_a^e, m_a^f]^T = A^{-1}B_A$  (11)

### 2.2.3 跟踪点选择算法

文献<sup>[12]</sup>指出，并非所有点都易于跟踪。文献<sup>[6]</sup>虽然也采用了基于仿射变换的刚体跟踪算法，但是由于没有考虑到刚体中各点的易跟踪性，所以其跟踪效果不好。易于跟踪的点 X 必须具备两个条件<sup>[12]</sup>：(1) 矩阵 (6) 的两个特征值较大。(2) 这两个特征值不能相差太大。但是同时具备这两个条件的点往往不多，如果采用这些点进行跟踪，就易受噪声干扰。本文注意到如果某点对应矩阵 (6) 的两个特征值中有一个较大，则此点在一个方向上有较好的跟踪特性，将这点加到整体跟踪模型中能提高整体在这个方向上的跟踪性能。所以本文成对选择分别具有不同较大特征值的点，使整体在每个方向上都有较好的跟踪性能。跟踪点选择算法如下：(1) 在视频对象区域中选择使得对应矩阵 (6) 的两个特征值都较大且相差不是很大的点。(2) 如果这样的点少于一定阈值，则在视频对象区域中成对选择分别具有不同较大特征值的点，直到没有这样的点对或是所选的点数达到一定阈值。

### 2.2.4 刚体分割算法

算法以当前帧  $F_t$  中视频对象的掩膜 ( $Mask_t$ ) 为输入，以下一帧  $F_{t+1}$  中视频对象掩膜 ( $Mask_{t+1}$ ) 为输出。

- (1) 初始整体平移  $M_T = [0,0]$ ，整体仿射  $M_A = [0,0,0,0,0,0]$
- (2) 用跟踪点选择算法在  $Mask_t$  中选择易于跟踪的跟踪点。
- (3) 用整体平移模型计算出  $Mask_t$  的整体平移变换向量  $M_T$ ,  $M_T = M_T + M_{T0}$ 。
- (4) 用  $M_T$  的逆变换对下一帧  $F_{t+1}$  进行平移



- (5) 若  $\| M_T \| > \text{Thresh}_T$  并且 平移变换的次数  $< \text{Iter}_T$  , 转 (3)
- (6) 如果此刚体只作平移运动, 转 (10)
- (7) 用整体仿射模型计算出  $\text{Mask}_k$  的整体仿射变换向量  $M_A$ ,  $M_A = M_A + M_{A_0}$ .
- (8) 用  $M_A$  的逆变换对下一帧  $F_{t+1}$  进行仿射变换
- (9) 若  $\| M_A \| > \text{Thresh}_A$  并且 仿射变换的次数  $< \text{Iter}_A$  , 转 (7)
- (10) 最终得到  $M_T = [m_t^a, m_t^d]^T$  ,  $M_A = [m_a^a, m_a^b, m_a^c, m_a^d, m_a^e, m_a^f]^T$  , 记  $M = [m_a^a + m_t^a, m_a^b, m_a^c, m_a^d + m_t^d, m_a^e, m_a^f]^T$  , 用  $M$  对  $\text{Mask}_k$  的边界点进行变换, 得到在下一帧中视频对象的边界点, 然后将边界点连成封闭边界线, 并区域填充算法 (见 2.2) 得到下一帧中的视频对象掩膜 ( $\text{Mask}_{k+1}$ ).

同时, 本文采用高斯层次模型不仅使算法适用于较大运动量, 并且增强抗噪能力。实现层次模型时, 在每一层中都运行上面的算法, 所不同的是初始的  $M_T$ 、 $M_A$  由上一层的结果乘以二给出, 最高层的  $M_T$ 、 $M_A$  被初始为零, 最低层的结果  $M_T$ 、 $M_A$  用来得到  $\text{Mask}_{k+1}$ 。

### 3 实验结果

#### 3.1 空域分割

基于 PBGUI 的空域分割结果如图 1 所示, 测试序列采用 CIF 格式的不同视频质量的 AKIYO 和 FOREMAN 序列。在图 1 (a) (b) 中, 左图上的实心白圆点表示用户在边界附近的输入点, 边界上的白线是 PBGUI 根据用户的输入点求出的轮廓线。右图是 PBGUI 根据用户输入点得到的视频对象。对于图像质量好, 边界清晰的 AKIYO 序列, 用户只需输入 9 个点, 而对于图像质量较差, 边界模糊的 FOREMAN 序列, 用户也只需输入 19 个点即可。

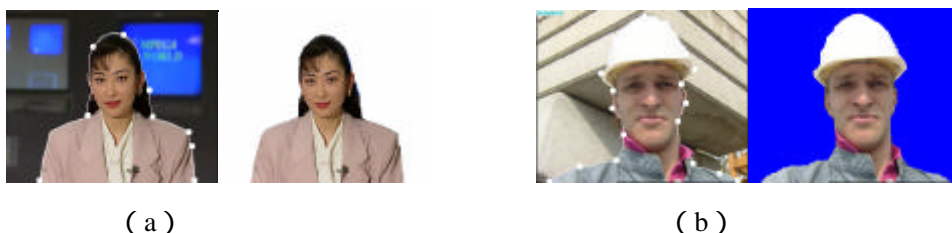


图 1 PBGUI 的空域分割结果

#### 3.2 基于边界跟踪的非刚体分割

基于边界跟踪的非刚体分割结果如图 2、图 3 所示, 测试序列同样采用 CIF 格式的不同视频质量的 AKIYO 和 FOREMAN 序列。第一帧分割采用空域分割方法, 结果如图 1 所示。对于 AKIYO 序列, 由于视频质量好, 同时运动量不大, 在整个分割过程中, 无任何用户干预, 分割效果相当好。



图 2 AKIYO 时域分割结果

对于 FOREMAN 序列, 由于视频质量较差, 视频对象的边界比较模糊, 同时运动量较大, 分割效果没有 AKIYO 序列好。图 3 所给出了没有经过任何用户调整的结果。由结果可以看出, 在 100 帧以

前分割效果是可以接受的。



图3 FOREMAN 时域分割结果

### 3.3 基于整体跟踪的刚体分割

基于整体跟踪的刚体分割结果如图4所示，测试序列采用 COASTGUARD 序列。在此序列中，选用 COASTGUARD 作为视频对象，作刚体运动。从 100 帧开始，COASTGUARD 完全独立出现，故本文从第 100 帧开始分割。图4中，第 100 帧是空域分割结果，以后的时域分割无任何用户干预，分割效果非常好。需要说明的是，分割效果的获得是由于采用了本文提出的跟踪点选择算法，因为 COASTGUARD 内部许多点少纹理，不易跟踪。如果采用文献<sup>[6]</sup>中的算法，将内部点全部用于跟踪，其效果如图5所示，在 120 帧时发生跟踪偏移，在 150 帧时偏移约 7 个像素，在 200 帧时偏移约 13 个像素，在 250 帧时偏移达 33 个像素，其效果已经不可接受了。



图4 COASTGUARD 时域分割结果

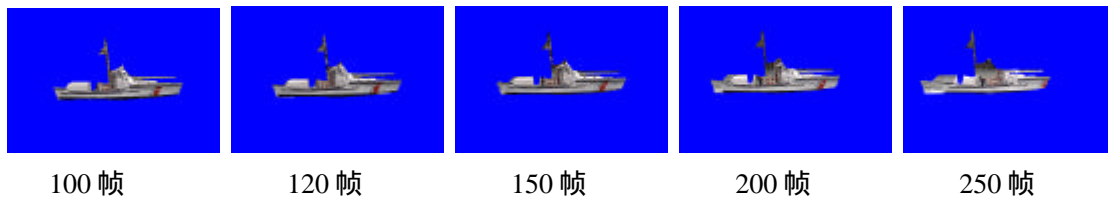


图5 采用文献[6]算法对 COASTGUARD 的分割结果

本文对刚体和非刚体的分割效果与文献<sup>[6]</sup>的效果做了分析对比，其结果见表6，

算法 \ 序列	AKIYO	FOREMAN	COASTGUARD
文献[6]算法	255 帧出现 8 像素以上偏移	70 帧出现 12 像素以上偏移	从 100 帧开始，在 210 帧出现 15 像素以上偏移
本文算法	对所有帧 (298 帧)，偏移小于 3 像素	100 帧出现 12 像素以上偏移	从 100 帧开始，到结束 (298 帧)，偏移小于 3 像素

## 4 结论

本文提出一种基于层次光流的半自动时空视频分割算法。算法由空域分割和时域分割组成。在空域分割中，本文提出一种基于点的图形用户界面(PBGUI)，可以精确地定义视频对象(VO)。时域分割对非刚体的边界跟踪和刚体的整体跟踪均采用了基于层次光流的算法，特别地，在刚体的整体跟踪时提出了跟踪点选择算法，极大地提高了分割性能。实验表明，本文的算法能够较精确地分割出视频对象，可用以实现面相对象编码、基于内容的多媒体应用、多媒体数据库的索引和检索等。在本文的算法中，还没有用到颜色信息，如何利用颜色信息来进一步改进算法性能将是本文的下一步工作。

## 参考文献：

- [1] MPEG Video and SNHC Groups. "Committee draft of MPEG-4, part 2, 14496-2" Technical Report ISO/IEC JTC/SC29/WG11/N1902, ISO/IEC[R], Fribourg, Switzerland, October 1997.
- [2] 季白杨,陈纯,钱英. 视频分割技术的发展[J]. 计算机研究与发展, 2001, 38(1):36-42.
- [3] Meier T, Ngan K N. Automatic segmentation of movings for video object plane generation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5):525-538
- [4] Wang D. Unsupervised video segmentation based on watersheds and temporal tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5):539-546
- [5] Gu Chuang, Lee Ming-Chieh. Semiautomatic segmentation and tracking of semantic video objects[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5):572-584
- [6] Kim Munchurl, Jeon J G, Kwak J S, Lee M H, Ahn C. Moving object segmentation in video sequences by user interaction and automatic object tracking[J]. Image and Vision computing, 2001, 19(5):245-260
- [7] Kass M, Witkin A, Terzopoulous D. Snakes: active contour models[J]. International Journal of Computer Vision,1988, 1(4):321-331
- [8] 章毓晋.图像分割[M].北京:科学出版社, 2001:34
- [9] Castleman K R. Digital Image Processing[M]. Englewood Cliffs, New Jersey: Prentice – Hall, 1996:460-461
- [10] Beauchemin S S, Barron J L, Fleet D J. On Optical flow[A]. Int. Conf. on Artificial Intelligence and Information-Control Systems of Robots [C]. Bratislava, Slovakia, 1994, 3-14
- [11] Lucas B D, Kanade T. An iterative image registration technique with and application to stereo vision[A]. Proceedings of the 7th International Joint Conference on Artificial Intelligence[C]. Vancouver, 1981,674-679
- [12] Shi J, Tomasi C. Good Features to Track[A].IEEE Conference on Computer Vision and Pattern Recognition[C]. Seattle, 1994, 593--600