

# 基于直方图的 Snake 视频对象跟踪算法

郭礼华 袁晓彤 李建华

(上海交通大学电子信息与电气工程学院, 上海 200030)

**摘要** 提出了一种基于直方图的 Snake 视频跟踪算法。该算法是基于相邻帧中对象在直方图信息上的一致性前提下,采用 Level Set 曲线演化算法来实现的。由于它在 Snake 能量函数中引入了全局统计特性——直方图信息,从而克服了过去采用均值、方差或运动矢量等局部特征的 Snake 跟踪算法中的局限性。通过此算法与基于局部特性的 Snake 跟踪算法的对比实验可以看出,该算法在实际跟踪效果方面有很好的改善。

**关键词** Snake 模型 主动轮廓模型 视频对象跟踪 直方图

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 1006-8961(2005)02-0197-06

## Video Object Tracking Method Based on Snake Model Using Object's Histogram Information

GUO Li-hua, YUAN Xiao-tong, LI Jian-hua

(School of Electronic & Information Technology, Shanghai JiaoTong University, Shanghai 200030)

**Abstract** This paper presents a video tracking algorithm based on Snake model using object's histogram information, which is designed on the basis of the object's histogram consistency between the consecutive image frames, and finally implemented by the use of the Level Set method. Because this algorithm uses the histogram information in the Snake energy function, it overcomes the limitation of the previous Snake video tracking algorithms based on the local characters, such as mean, deviation or motion vector. The contrastive experiment between our algorithm and those based on the local characters reveals that our algorithm performs better than those based on local characters.

**Keywords** Snake model, active contour model, video object tracking, histogram

## 1 引言

视频对象跟踪技术是计算机视觉领域中的一项目基础性,也是非常重要的研究课题。它和视频对象分割技术组成了基于对象视频处理的基础性研究课题。视频跟踪技术中的一些算法和研究成果现已很好地应用于基于对象的视频压缩、数字化可视操作、基于对象的视频检索和对象交互式媒体操作等领域中。

目前已经提出了很多的视频对象跟踪技术,其中早期的跟踪算法主要是关注于特征点<sup>[1]</sup>和边界

点的跟踪技术<sup>[2]</sup>。由于跟踪算法的最终目标是实现整个区域跟踪,而不是单对其中特征点或边界点跟踪,因此在这些早期跟踪算法中不得不提出了很多的关于区域形状和区域运动的假设性条件。为了克服基于特征点的区域跟踪算法的这些缺陷,又提出了基于主动轮廓模型算法<sup>[3]</sup>和主动网格模型算法<sup>[4]</sup>。主动轮廓模型也称为 Snake 模型,它主要是定义了 Snake 积分能量函数,其能量函数包括内部能量、图像能量和外部能量。对象跟踪的依据就是使得其所有的能量函数积分最小。Kass 等人<sup>[3]</sup>还给出了利用变分法来求解积分最小的方法。由于变

**基金项目:** 国家“863”高科技项目(2002AA145090)

**收稿日期:** 2003-11-18; **改回日期:** 2004-06-25

**第一作者简介:** 郭礼华(1978~),男,现为上海交通大学电子信息与电气工程学院博士研究生。研究方向包括图像分割、运动对象跟踪、数字水印和视频内容分析等领域。E-mail: guolihua@stju.edu.cn

分法在实际的求解过程中,数值解不稳定,并且在复杂的图像拓扑结构中不能很好地进行曲线演化,为此提出了一些改进模型,如离散的 Snake 模型<sup>[5]</sup>和 Geodesic 主动轮廓模型<sup>[6]</sup>等。离散的 Snake 模型采用动态编程的方法寻找最小点,而 Geodesic 主动轮廓模型采用 Level Set 方法<sup>[7]</sup>作为曲线演化算法。它把曲线演化过程模拟成一个曲面演化过程。根据拓扑的改变自适应地演变曲线,并且它的作用范围不再局限于曲线上的点,而是整个图像平面域。这种方法很好地解决了采用变分法进行数值求解所产生的问题,并且其对初始曲线的选择也不敏感。Level Set 方法由于这些特有的优势,现已广泛地被各种 Snake 模型作为其曲线演化算法。

过去的 Snake 跟踪算法,主要是利用跟踪区域中的局部特征一致性条件,采用 Level Set 曲线演化算法来实现 Snake 跟踪算法。如考虑到视频对象在运动过程中,其区域的颜色均值和方差局部特征基本保持不变这个特点,采用 Level Set 方法,实现了一种基于区域信息的 Snake 跟踪模型<sup>[8]</sup>;也是采用 Level Set 方法,但是考虑的不是区域的颜色均值和方差局部特征,而是利用区域运动中运动矢量,建立了一种不需要运动计算的 Snake 跟踪模型<sup>[9]</sup>;结合区域均值信息和边缘梯度信息,用 Level Set 方法进行运动目标检测和跟踪<sup>[10]</sup>。这些跟踪算法中,Snake 能量函数采用的都是局部统计特性,其应用的局限性是非常明显的。为此,提出一种基于直方图的 Snake 视频跟踪算法。这种算法是基于相邻帧中对象在直方图信息上的一致性前提条件下,采用 Level Set 曲线演化算法来实现的。由于它在 Snake 能量函数中引入了全局统计特性——直方图信息,从而克服了过去采用均值、方差或运动矢量等局部特征的 Snake 跟踪算法的局限性。这种基于直方图全局统计特性的 Snake 跟踪算法在实际的跟踪效果上得到很好的改善。

## 2 模型描述

在视频对象跟踪过程中,一般可把当前帧  $I_n$  分成两个区域,一个区域对应所跟踪的视频对象,用数学符号  $\Omega_n^{\text{in}}(x, y)$  表示;另一区域对应于背景,用数学符号  $\Omega_n^{\text{out}}(x, y)$  来表示。而两个区域之间的边界曲线用  $L_n(x, y)$  来表示。Snake 模型首先提取视频对象区域、背景区域和边界曲线之间的纹理、运动和

统计等特性,建立相应的特征函数  $J_n(\Omega_n^{\text{in}}(x, y), \Omega_n^{\text{out}}(x, y), L_n(x, y))$ 。接着利用所建立的特征函数,定义 Snake 能量函数

$$E = \iint_{I_n} J_n(\Omega_n^{\text{in}}(x, y), \Omega_n^{\text{out}}(x, y), L_n(x, y)) dx dy$$

最后在 Snake 能量函数  $E$  最小化准则下,Snake 曲线通过曲线迭代演化,使得 Snake 曲线收敛于实际的对象边界,从而最终实现视频对象的分割与跟踪。

**假设 1** 对象在移动的过程中,其跟踪对象的统计特性在相邻的两帧内基本保持不变,也就是其跟踪对象的直方图信息在相邻的两帧内基本保持不变。

**假设 2** 对象在移动的过程中,其背景的统计特性在相邻的两帧内基本保持不变,也就是其背景的直方图信息在相邻的两帧内基本保持不变。

基于以上的两个假设条件,定义视频对象跟踪准则。

**准则 1** 对象在移动的过程中,当前帧内曲线所围成的区域内部直方图和前一帧曲线所围区域内部直方图差别最小,同时其相应的外部直方图差别也要最小。

根据以上视频对象跟踪准则,Snake 能量函数可表示为

$$E = \iint_{I_n} J_n(\Omega_n^{\text{in}}(x, y), \Omega_n^{\text{out}}(x, y), L_n(x, y)) dx dy \\ = \iint_{I_n} \Psi[H(\Omega_n^{\text{in}}(x, y)), H(\Omega_{n-1}^{\text{in}}(x, y))] dx dy + \\ \iint_{I_n} \Psi[H(\Omega_n^{\text{out}}(x, y)), H(\Omega_{n-1}^{\text{out}}(x, y))] dx dy \quad (1)$$

$H(\Omega_n^{\text{in}}(x, y)), H(\Omega_n^{\text{out}}(x, y))$  分别是当前帧内曲线所围的内部区域和外部区域的直方图,  $H(\Omega_{n-1}^{\text{in}}(x, y)), H(\Omega_{n-1}^{\text{out}}(x, y))$  分别是前一帧内曲线所围的内部区域和外部区域的直方图。 $\Psi(\cdot)$  函数表示两个直方图的偏差函数。在跟踪算法中,直方图  $H$  可分成  $N$  等级,对两个直方图中的每个等级所对应的偏离度进行求和,这个求和值便是两个直方图值的偏差函数。区域  $a$  和区域  $b$  之间的直方图偏差函数表示如下:

$$\Psi(H(a), H(b)) \\ = \sum_{i=1}^N \left( \frac{H_i(a) - H_i(b)}{\min(\sum_{i=1}^N H_i(a), \sum_{i=1}^N H_i(b))} \right) \quad (2)$$

其中,  $H_i(\cdot)$  表示区域第  $i$  等级的直方图值。由于采用了式(2)的偏差函数,所以 Snake 能量函数又可表示为

$$E = \iint_h \sum_{i=1}^N \left( \frac{H_i(\Omega_n^{in}(x,y)) - H_i(\Omega_{n-1}^{in}(x,y))}{\min(\sum_{i=1}^N H_i(\Omega_n^{in}(x,y)), \sum_{i=1}^N H_i(\Omega_{n-1}^{in}(x,y)))} \right) dx dy + \iint_h \sum_{i=1}^N \left( \frac{H_i(\Omega_n^{out}(x,y)) - H_i(\Omega_{n-1}^{out}(x,y))}{\min(\sum_{i=1}^N H_i(\Omega_n^{out}(x,y)), \sum_{i=1}^N H_i(\Omega_{n-1}^{out}(x,y)))} \right) dx dy \quad (3)$$

在相邻的视频帧内,由于对象区域的直方图信息变化不大,所以式(3)中的  $\min(\sum_{i=1}^N H_i(\Omega_n^{in}(x,y)), \sum_{i=1}^N H_i(\Omega_{n-1}^{in}(x,y)))$  可以用  $\sum_{i=1}^N H_i(\Omega_{n-1}^{in}(x,y))$  代替,  $\min(\sum_{i=1}^N H_i(\Omega_n^{out}(x,y)), \sum_{i=1}^N H_i(\Omega_{n-1}^{out}(x,y)))$  可以用  $\sum_{i=1}^N H_i(\Omega_{n-1}^{out}(x,y))$  代替。从而式(3)可以简化成

$$E = \iint_h \sum_{i=1}^N \left( \frac{H_i(\Omega_n^{in}(x,y)) - H_i(\Omega_{n-1}^{in}(x,y))}{\sum_{i=1}^N H_i(\Omega_{n-1}^{in}(x,y))} \right) dx dy + \iint_h \sum_{i=1}^N \left( \frac{H_i(\Omega_n^{out}(x,y)) - H_i(\Omega_{n-1}^{out}(x,y))}{\sum_{i=1}^N H_i(\Omega_{n-1}^{out}(x,y))} \right) dx dy \quad (4)$$

在此,引入 Dirac 函数  $\delta(\cdot)$ , 则  $H_i(\Omega_n^{in}(x,y)) - H_i(\Omega_{n-1}^{in}(x,y))$  用 Dirac 函数简化成

$$H_i(\Omega_{n-1}^{in}(x,y)) \delta(H_i(\Omega_{n-1}^{in}(x,y)))$$

$H_i(\Omega_n^{out}(x,y)) - H_i(\Omega_{n-1}^{out}(x,y))$  用 Dirac 函数简化成

$$H_i(\Omega_{n-1}^{out}(x,y)) \delta(H_i(\Omega_{n-1}^{out}(x,y)))$$

此时式(4)简化为

$$E = \iint_h \left( \frac{H_i(\Omega_{n-1}^{in}(x,y))}{\sum_{i=1}^N H_i(\Omega_{n-1}^{in}(x,y))} \right) \times \delta(H_i(\Omega_{n-1}^{in}(x,y))) dx dy + \iint_h \left( \frac{H_i(\Omega_{n-1}^{out}(x,y))}{\sum_{i=1}^N H_i(\Omega_{n-1}^{out}(x,y))} \right) \times \delta(H_i(\Omega_{n-1}^{out}(x,y))) dx dy \quad (5)$$

在式(5)中,  $\frac{H_i(\Omega_{n-1}^{in}(x,y))}{\sum_{i=1}^N H_i(\Omega_{n-1}^{in}(x,y))}$  其实就是参考帧内

曲线所围的内部区域的直方图中  $i$  等级像素在整个直方图中所占的比例,  $\frac{H_i(\Omega_{n-1}^{out}(x,y))}{\sum_{i=1}^N H_i(\Omega_{n-1}^{out}(x,y))}$  是参考帧

内曲线所围的外部区域的直方图中  $i$  等级像素在整个直方图中所占的比例, 分别用  $K_i(\Omega_{n-1}^{in}(x,y))$ ,  $K_i(\Omega_{n-1}^{out}(x,y))$  比例函数来表示。式(5)可进一步简化成

$$E = \iint_h K_i(\Omega_{n-1}^{in}(x,y)) \delta(H_i(\Omega_{n-1}^{in}(x,y))) dx dy + \iint_h K_i(\Omega_{n-1}^{out}(x,y)) \delta(H_i(\Omega_{n-1}^{out}(x,y))) dx dy \quad (6)$$

由于视频对象区域  $\Omega_{n-1}^{in}(x,y)$  和背景区域  $\Omega_{n-1}^{out}(x,y)$  是互补的, 所以

$$\delta(H_i(\Omega_{n-1}^{out}(x,y))) = -\delta(H_i(\Omega_{n-1}^{in}(x,y)))$$

Snake 能量函数为

$$E = \iint_h K_i(\Omega_{n-1}^{in}(x,y)) \delta(H_i(\Omega_{n-1}^{in}(x,y))) dx dy - \iint_h K_i(\Omega_{n-1}^{out}(x,y)) \delta(H_i(\Omega_{n-1}^{in}(x,y))) dx dy = \iint_h [K_i(\Omega_{n-1}^{in}(x,y)) - K_i(\Omega_{n-1}^{out}(x,y))] \times \delta(H_i(\Omega_{n-1}^{in}(x,y))) dx dy \quad (7)$$

根据 Euler-Lagrange 方程可得

$$\frac{\delta E}{\delta L_n} = [-K_i(\Omega_{n-1}^{in}(x,y)) + K_i(\Omega_{n-1}^{out}(x,y))] \mathbf{n}$$

其中,  $\mathbf{n}$  是曲线  $L$  的单位外法向量。具体推导过程请参见文献[7]、[11]。能量函数取极值的必要条件是  $\frac{\delta E}{\delta L_n} = 0$ , 为了得到轮廓曲线  $L_n$  的变形方程, 为  $L_n$  加上一个时间变量  $t$ , 记为  $L_n(t)$ , 其中,  $t$  为曲线函数的时间参数。根据文献[7], 曲线函数  $L_n(t)$  与 Snake 能量函数之间有以下变形关系:

$$\frac{\partial L_n}{\partial t} = -\frac{\delta E}{\delta L_n} \quad (8)$$

从而  $\frac{\partial L_n}{\partial t} = [K_i(\Omega_{n-1}^{in}(x,y)) - K_i(\Omega_{n-1}^{out}(x,y))] \mathbf{n}$ , 如果给定初始轮廓是  $L_0$ , 就得到了下面的偏微分方程

$$\begin{cases} \frac{\partial L_n}{\partial t} = [K_i(\Omega_{n-1}^{in}(x,y)) - K_i(\Omega_{n-1}^{out}(x,y))] \mathbf{n} \\ L_n(0) = L_0 \end{cases} \quad (9)$$

根据这个方程, 从初始轮廓  $L_0$  出发, 当  $t$  充分大时

$L_n$  收敛于局部极值,也就是  $\frac{\delta E}{\delta L_n} = -\frac{\partial L_n}{\partial t} = 0$ ,这时得到的  $L_n$  就是目标轮廓。在实际的 Snake 模型实现中,对 Snake 曲线增加一个额外的附加内力  $\mu\kappa n$ ,该力与曲率成正比,其作用的结果是使轮廓曲线变得光滑。实际的偏微分方程为

$$\begin{cases} \frac{\partial L_n}{\partial t} = [\mu\kappa + K_i(\Omega_{n-1}^{\text{in}}(x,y)) - K_i(\Omega_{n-1}^{\text{out}}(x,y))]n \\ L_n(0) = L_0 \end{cases} \quad (10)$$

方程(10)有比较合理的力学解释,把轮廓  $L$  的动态变形看成是作用在其上的两种力的合力下的运动。第 1 种力是控制轮廓光滑程度的“内力” $\mu\kappa n$ ,该力与曲率成正比,其作用的结果是使轮廓曲线变得光滑;第 2 种力是控制轮廓运动趋势的“外力” $(K_i(\Omega_{n-1}^{\text{in}}(x,y)) - K_i(\Omega_{n-1}^{\text{out}}(x,y)))n$ ,若当前轮廓上的点  $P(x,y)$  属于背景,则此点的像素值对背景的直方图贡献应该远大于对前景对象直方图的贡献。

此时该点处像素值对应于背景直方图第  $i$  等级处的比例明显要大于其对应于前景直方图第  $i$  等级处的比例。即  $K_i(\Omega_{n-1}^{\text{out}}(x,y)) \gg K_i(\Omega_{n-1}^{\text{in}}(x,y))$ ,从而使得  $K_i(\Omega_{n-1}^{\text{in}}(x,y)) - K_i(\Omega_{n-1}^{\text{out}}(x,y)) < 0$ ,由于此时的外力作用为负值,曲线将沿着曲线的法线反方向进行收缩,对应于图 1(b)。若当前轮廓上的点  $P(x,y)$  属于前景,则此点的像素值对前景对象的直方图贡献应该远大于对背景直方图的贡献。此时该点处像素值对应于前景直方图第  $i$  等级处的比例明显要大于其对应于背景直方图第  $i$  等级处的比例,即  $K_i(\Omega_{n-1}^{\text{out}}(x,y)) \ll K_i(\Omega_{n-1}^{\text{in}}(x,y))$ ,从而使得  $K_i(\Omega_{n-1}^{\text{in}}(x,y)) - K_i(\Omega_{n-1}^{\text{out}}(x,y)) > 0$ ,由于此时的外力作用为正值,曲线将沿着曲线的法线方向进行扩展,对应于图 1(c)。当轮廓变形到前景的轮廓附近时,轮廓的变化将进入一种动态平衡,即交替运动到前景和背景区域,这时的轮廓就是目标轮廓。

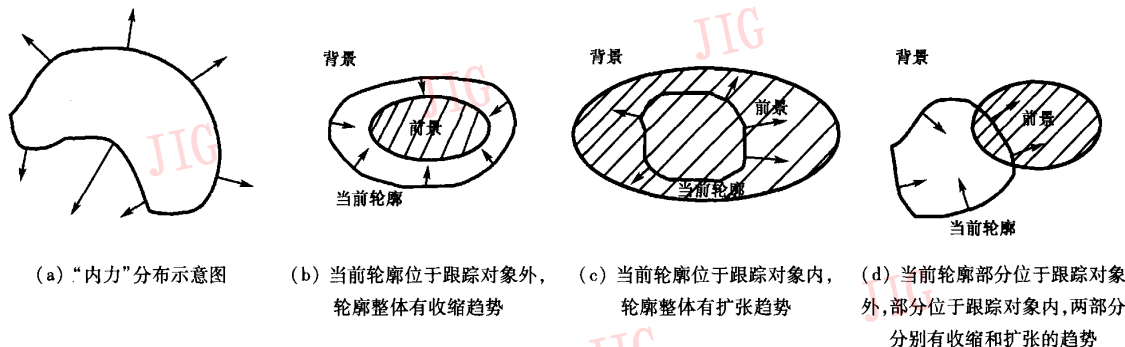


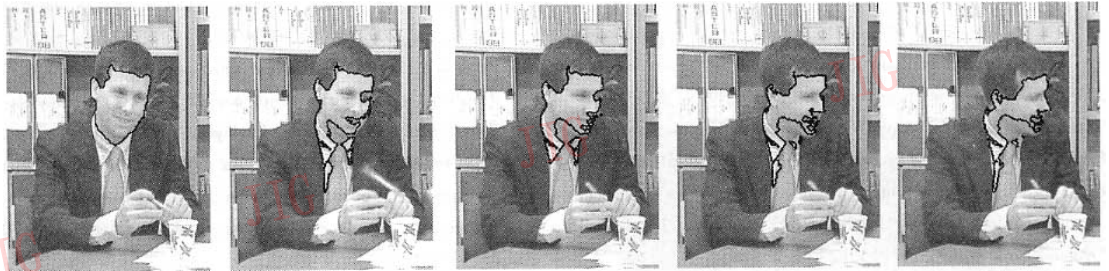
图 1 各种情况的 Snake 曲线演变示意图  
Fig.1 The curve evolution in different situations

曲线演化方程(10)的实现采用 Level Set 方法,具体数值解参见文献[12]。

### 3 实验结果与讨论

比较基于直方图的 Snake 跟踪算法和文献[7]、[8]所提及的跟踪算法的跟踪效果。文献[8]采用的是区域信息的局部均值特征,而文献[9]采用的是运动矢量特征。本文的跟踪算法和文献[8]、[9]所提及的跟踪算法都是采用 Level Set 方法作为其曲线演化算法。选取 MPEG 标准测试序列作为测试源,跟踪结果如图 2 所示。在图 2 中,其最初始的

跟踪曲线都是人为地给出的。从跟踪结果可以看出,文献[8]采用的均值方法在亮度分布比较集中和背景亮度反差比较大的时候,其跟踪效果比较好,可是当区域亮度和背景亮度反差小的时候,其跟踪效果干扰比较大。文献[9]采用的是运动矢量一致性特性。由于在对对象的运动矢量求解过程中存在运动暴露问题,使得因为对象移动所引起的暴露部分不能准确估计其相应的运动矢量,所以其跟踪效果中存在明显的“拖尾效应”。从这个对比实验可以看出,采用了全局的统计直方图信息的 Snake 跟踪算法后,其实际跟踪效果比采用局部统计特性的 Snake 跟踪算法有很大的改善。



(a) 文献[8]方法



(b) 文献[9]方法



(c) 本文方法

图2 文献[8]、[9]及本文的 Snake 跟踪算法的跟踪效果对比图

Fig. 2 The comparison results between the our method with the tracking methods mentioned in the paper[Gast02, Man02]

## 4 结论

提出了一种基于直方图的 Snake 视频跟踪算法。由于这种算法在 Snake 能量函数中引入了全局统计特性——直方图信息,从而克服了过去采用均值、方差或运动矢量等局部特征的 Snake 跟踪算法中的局限性。从这种算法与基于局部特性的 Snake 跟踪算法的对比实验结果可以看出,本文算法在实际跟踪效果方面有很好的改善。虽然该算法在视频对象跟踪方面有很好的改善,但是它也有其局限性。首先,当区域的直方图分布和周围区域的直方图分布比较接近时,算法有可能会陷入局部最小;其次就是由于 Snake 模型主要的缺点是算法复杂,运算时间长,如果应用到实时的视频对象跟踪中还有一定

的距离。所以下一步的研究目标是如何进一步改善 Snake 跟踪模型的有效性和其运算速度,同时研究快速 Snake 跟踪算法。

## 参考文献 (Reference)

- 1 Sethi I K, Jain R. Finding trajectories of feature points in a monocular image sequence [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(1): 56 ~ 73.
- 2 Crowley J L, Stelmazyk P, Discours C. Measuring image flow by tracking edge lines [A]. In: Proceedings of IEEE Computer Vision and Pattern Recognition [C], San Diego, California, 1989: 658 ~ 664.
- 3 Kass M, Witkin A, Terzopoulos D. Snakes: active contour models [A]. In: Proceedings of First International Conference on Computer Vision [C], London, England, 1987: 259 ~ 268.
- 4 Valette S, Magnin I, Prost R. Active mesh for video segmentation and objects tracking [A]. In: International Conference on Image

- Processing[C], Thessaloniki, Greece, 2001:77 ~ 80.
- 5 Amini A A, Weymouth T E, Jain R C. Using dynamic programming for solving variational problems in vision[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990,12(9):855 ~ 867.
  - 6 Caselles V, Kimel R, Sapiro G. Geodesic active contours [J]. International Journal of Computer Vision, 1997,22(1):61 ~ 79.
  - 7 Osher S, Sethian J A. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation[J]. Journal Computer Physics, 1988,79(1):12 ~ 49.
  - 8 Gastaud M, Barlaud M, Aubert G. Tracking video objects using active contours[A]. In: Proceedings of the Workshop on Motion and Video Computing[C], Orlando, Florida, 2002:90 ~ 95.
  - 9 Mansouri A R. Region tracking via level set PDEs without motion computation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002,24(7):947 ~ 961.
  - 10 Paragios N, Deriche R. Geodesic active contours and level sets for the detection and tracking of moving objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000,22(3):266 ~ 280.
  - 11 Jehan-Besson S, Barlaud M. Dream<sup>2</sup>s: deformable regions driven by an eulerian accurate minimization method for image and video segmentation[J]. International Journal of Computer Vision, 2003, 53(1):45 ~ 70.
  - 12 Goldenberg R, Kimmel R, Rivlin E, *et al.* Fast geodesic active contours [J]. IEEE Transactions on Image Processing, 2001, 10(10):1467 ~ 1475.