

基于区域活动轮廓运动目标跟踪方法研究

查宇飞 张育 毕笃彦

(空军工程大学工程学院信号与信息处理实验室, 西安 710038)

摘要 根据贝叶斯估计理论,首先建立了图像序列中运动目标的跟踪模型,然后用高斯分布来描述图像的区域信息,并通过对模型的分析,与区域活动轮廓模型建立对应关系,将问题的求解转化为能量最小化问题。同时为了克服目标在运动中发生的拓扑形变,采用水平集方法进行数值实现。实验结果表明,这种方法不仅可以对多个运动目标进行跟踪,并能非常好地逼近运动目标的轮廓,而且能够自然地处理运动目标的拓扑形变。

关键词 活动轮廓 运动目标跟踪 水平集

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2006)12-1844-05

Moving Objects Tracking Using Region-based Active Contours

ZHA Yu-fei, ZHANG Yu, BI Du-yan

(Signal & Information Processing Laboratory, Engineering College of Air Force Engineering University, Xi'an 710038)

Abstract Firstly, tracking model of moving objects in images is established based on Bayesian theory. Secondly, image regional information is described by Gaussian distribution. At the same time, the tracking model is matched to region-based active contours through analyzing, and then this problem is converted into the minimizing of energy. Lastly, level set is used to solve the PDE equation, which can process topology naturally. The result shows that the proposed method not only could track multi-objects and approach to the objects contours, but also does well when the moving objects deforms topologically.

Keywords active contours, moving objects tracking, level set

1 引言

在图像序列中,对“敏感”运动目标进行跟踪在监视、视频编码和计算机视觉等领域有着广泛的应用前景,但到目前为止它仍然是一个比较困难的问题。活动轮廓模型(active contour models)分为参数活动轮廓模型^[1]和几何活动轮廓模型。参数活动轮廓直接用曲线或曲面参数化形式来表示曲线或曲面的变形。基本思想是通过最小化能量函数来实现对图像的分割和跟踪,最小化的能量函数可以写成如下形式:

$$E(C) = \int_0^l (E_{\text{internal}}(v(s)) + E_{\text{image}}(v(s))) ds \quad (1)$$

其中, s 是轮廓曲线 C 的弧长, E_{internal} 表示曲线弯曲所产生的内部能量, E_{image} 表示在图像中的力。这种方法的自由度大,可以逼近任何形状不规则的曲线,并可约束模型具有目标的先验轮廓信息。但这种模型难以处理拓扑结构的变化,如曲线的合并和分裂。几何活动轮廓模型可以很自然地处理拓扑结构的变化,这种模型主要是基于曲线演化(curve evolution)理论和水平集(level set)理论。几何活动轮廓模型的基本方程为

$$\frac{\partial C}{\partial t} = \mathcal{K}(\kappa, I)n \quad (2)$$

其中, $\mathcal{K}(\kappa, I)$ 是与图像 I 和曲线曲率 κ 相关的速度, n 是曲线的单位法向量。由上式可见,曲线 C 以速度 V 沿着曲线法向运动。在速度 V 的函数中,曲率

κ 的作用类似于参数模型中的内力,保持曲线 C 的光滑性,而与图像有关的项相当于外力。

图像力一般用梯度图像 ∇I 来表示,后来 Caselles^[2] 等人用函数 $g(\nabla I)$ 来表示,其中 $g(\cdot)$ 是 S 形函数。对于纹理图像和含有较大噪声的图像,就不能直接从梯度图像 ∇I 中来分辨目标的边界,当曲线在其指引下变形时,很容易陷入局部极小,收敛到错误的位置。为了克服这种缺陷,Zhu^[3] 等人提出了区域竞争的方法,利用区域的信息——这一全局量,避免陷入局部极小错误。在文献[4]中,Jehan-Besson 把它称作区域描述符。Paragios^[5] 等人在几何活动轮廓中加入了这种能量项,进行检测和跟踪。

本文根据贝叶斯估计理论,建立图像序列中运动目标跟踪的模型,然后对建立的模型进行分析,巧妙地与基于区域的活动轮廓模型建立对应关系,将运动跟踪问题的求解转化为能量最小化的问题。为了克服运动目标在运动中发生拓扑形变,采用水平集方法进行数值实现。实验结果表明,这种方法不仅可以跟踪多个运动目标且能非常好地逼近运动目标的轮廓,消除运动目标的拓扑形变的影响。

2 基于区域的活动轮廓

目标跟踪的方法大致可以分为基于区域的跟踪和基于边缘的跟踪。基于区域的跟踪是利用目标的区域的信息,如灰度直方图等。这种方法可以很好地跟踪目标,但是不能精确地捕捉目标边缘,而基于边缘的跟踪是利用目标的边缘信息,因此可精确捕捉目标边缘,但是这种方法依赖于目标清晰的轮廓,但在实际的图像中,目标的轮廓往往是模糊的,因此这种方法不能理想地跟踪目标,甚至会丢失目标。基于以上的原因,现行的跟踪方法大部分都是将目标的区域信息和边缘信息相结合,从而更好地进行跟踪。

活动轮廓模型是一种基于目标边缘信息方法,Paragios^[5] 结合 Zhu^[3] 的区域竞争思想,提出了测地活动区域的方法,这种方法首先用统计理论得到不同区域及其相关边界的概率,然后根据概率得到曲线演化方程:

$$\frac{\partial C}{\partial t} = \alpha \log \left(\frac{P_B(K(C))}{P_A(K(C))} \right) n + (1 - \alpha) \nabla g(P_{\text{con}}(K(C))) \kappa - \nabla g(P_{\text{con}}(K(C))) n \quad (3)$$

其中 P_A, P_B 分别表示区域同质属性分布的条件概率密度函数, P_{con} 表示图像中任意点位于图像实际边界的概率。这种模型充分利用了区域和边缘的信息,取得了较好的效果,但当图像中具有多个目标时,需要多个模型的耦合完成,计算复杂。

Chen 和 Vese^[6] 简化 Mumford-Shad^[7] 模型,得到如下的能量函数:

$$E(C) = \int_{\text{inside}} |K(x, y) - c_1|^2 dx dy + \int_{\text{outside}} |K(x, y) - c_2|^2 dx dy \quad (4)$$

其中 c_1, c_2 分别是图像 I 被曲线 C 所分隔的内部图像和外部图像的平均灰度值。这种模型在噪声较大,甚至是不存在实际意义的边界时仍然能够有效地对图像进行分割。

Jehan-Besson^[4] 等人提出了基于区域的活动轮廓的更一般的模型:

$$E(C) = \int_{\text{inside}} k^{\text{inside}}(C) dx dy + \int_{\text{outside}} k^{\text{outside}}(C) dx dy + \int_C k^{\text{contour}}(C) ds \quad (5)$$

其中 $k^{\text{contour}}(\cdot), k^{\text{inside}}(\cdot), k^{\text{outside}}(\cdot)$ 分别表示曲线 C 所在边缘的函数,以及被曲线 C 所分隔的内部和外部区域的函数。

3 目标跟踪模型

首先从贝叶斯理论出发,建立序列图像中运动目标的跟踪模型,然后将这种模型与基于区域的活动轮廓模型建立对应联系。

3.1 基于贝叶斯估计的目标跟踪模型

设封闭曲线 C 为图像 I 中运动目标的轮廓,则可将目标所在的区域表示为 $R_{\text{obj}} \in \Omega$,背景区域表示为 $R_{\text{bak}} \in \Omega$,其中 $\Omega \in R^2$ 。它们与图像 I 的关系是:

$$I = R_{\text{obj}} \cup R_{\text{bak}} \cup C \quad (6)$$

若在第 $k(k \in N)$ 帧图像 I^k 中,运动目标的轮廓为 C_{origin} ,跟踪目的就是要在 $k+1$ 帧图像 I^{k+1} 中,寻找运动目标的新轮廓 C_{new} ,即对 C_{new} 进行最佳估计。

由于本文所研究的是简单背景情况,模型中不考虑目标被遮挡或覆盖的情况,因此观测模型可以表述为

$$I^{k+1}(\varphi(C)) = I^k(C) + \nu \quad (7)$$

其中 $\varphi(C_{\text{origin}}) = C_{\text{new}}, \varphi: \Omega \rightarrow \Omega, \nu$ 是系统的随机噪声,服从均值为 0,协方差矩阵为 Σ 的 2 维高斯

分布。

运动目标跟踪 就是在已知 k 帧图像 I^k 中目标的轮廓 C^k 的情况下,在第 $k+1$ 帧图像 I^{k+1} 中对运动目标新的轮廓 C^{k+1} 进行最佳估计 \hat{C}^{k+1} 。采用最大后验概率估计(MAP),使概率 $P(C^{k+1} = C_{true}^k | I^{k+1}, I^k, C^k)$ 最大的 C^{k+1} ,也就是

$$\hat{C}^{k+1} = \arg \max C(C^{k+1} = C_{true}^{k+1} | I^{k+1}, I^k, C^k) \quad (8)$$

根据贝叶斯公式,可以进一步得到:

$$\hat{C}^{k+1} = \arg \max P(I^{k+1} | I^k, C^k, C^{k+1} = C_{true}^{k+1}) P(C^{k+1} = C_{true}^{k+1} | I^k, C^k) \quad (9)$$

又由于 $J = R_{obj} \cup R_{bak} \cup C$,因此

$$P(I^{k+1}(X) | I^k, C^k, C^{k+1} = C_{true}^{k+1}) = \begin{cases} P_{obj}(I^{k+1}(X) | I^k, C^k, C^{k+1} = C_{true}^{k+1}) & X \in R_{obj} \\ 0 & X \in C \\ P_{bak}(I^{k+1}(X) | I^k, C^k, C^{k+1} = C_{true}^{k+1}) & X \in R_{bak} \end{cases} \quad (10)$$

其中 X 为一向量,表示图像 I^{k+1} 中像素所在的位置。将式(10)带入式(9)得:

$$\hat{C}^{k+1} = \arg \max_{R_{obj} R_{bak} \in \Omega} \left\{ \left(\prod_{X \in R_{obj}} P_{bak}(I(X) | I^k, C^k, C^{k+1} = C_{true}^{k+1}) \right) \times \left(\prod_{X \in R_{bak}} P_{bak}(I(X) | I^k, C^k, C^{k+1} = C_{true}^{k+1}) \right) \times P(C^{k+1} = C_{true}^{k+1} | I^k, C^k) \right\} \quad (11)$$

3.2 与区域活动轮廓的关系

取概率 P_{obj}, P_{bak} 的负对数作为其所对应的能量;由于梯度图像是图像的隐含信息(梯度图像可以通过图像得到),因此,对于式(11)最后一项

$$P(C^{k+1} = C_{true}^{k+1} | I^k, C^k) = P(C^{k+1} = C_{true}^{k+1} | I^k, \nabla I^{k+1}, C^k) \quad (12)$$

可看作是对运动目标轮廓形状的约束,取其负对数作为目标轮廓形状和图像梯度力所对应的能量,即

$$E_{shape} = -\log P(C^{k+1} = C_{true}^{k+1} | I^k, \nabla I^{k+1}, C^k) = \mu \int_C ds + \nu \int_C g(\|\nabla I^{k+1}\|) ds \quad (13)$$

其中 μ 是常数, s 表示弧长, $g(x) \propto \exp\left(-\frac{x^2}{2\sigma_g^2}\right)$ 为高斯函数。于是将最大后验概率估计的问题转化为能量最小化的问题。即

$$\hat{C}^{k+1} = \arg \min E(C^{k+1} = C_{true}^{k+1} | I^k, I^{k+1}, C^k) \quad (14)$$

因此能量函数可以写作:

$$E = E_{in} + E_{out} + E_{shape} = -\int_{R_{obj}} \log(P_{obj}(I^{k+1}(X) | I^k, C^k, C^{k+1} = C_{true}^{k+1})) dx dy -$$

$$\int_{R_{bak}} \log(P_{bak}(I^{k+1}(X) | I^k, C^k, C^{k+1} = C_{true}^{k+1})) dx dy - \mu \int_C ds - \nu \int_C g(\|\nabla I^{k+1}\|) ds \quad (15)$$

要使上式的能量最小化,根据文献[3],解 Euler-Lagrange 方程,可得到:

$$\frac{dC}{dt} = [\log P_{obj}(I^{k+1} | I^k, C^k, C^{k+1} = C_{true}^{k+1}) - \log P_{bak}(I^{k+1} | I^k, C^k, C^{k+1} = C_{true}^{k+1}) - \mu \kappa - \nu (g(\|\nabla I^{k+1}\|) \kappa - \nabla g(\|\nabla I^{k+1}\|))] n \quad (16)$$

其中 κ 是轮廓 C 的曲率 $\kappa = \frac{\ddot{x}\dot{y} - \dot{x}\ddot{y}}{(\dot{x} + \dot{y})^{3/2}}$ ($\dot{x}, \dot{y}, \ddot{x}, \ddot{y}$) 表示轮廓 C 各点的一阶微分和二阶微分, n 表示轮廓 C 指向外的单位法向量。概率 $P_{\alpha}, \alpha \in \{obj, bak\}$ 的计算,文献[8]给出如下:

$$P_{\alpha} = C_{\alpha} \sup_{\{Z: \|Z\| < \delta, X+Z \in R_{\alpha}\}} P_{\alpha}, \alpha \in \{obj, bak\} \quad (17)$$

它是对图像中的每一个像素和它周围的像素进行比较,取其中概率最大者来判断该像素是属于背景还是目标。但是,这种方法的计算量非常大,而且当范围不合适的时候,会进行错误的跟踪。针对这些不足,与文献[8]不同,本文是利用区域的全局信息,轮廓 C 在全局的信息的指引下形变,避免了陷入局部极小,并且提高了计算效率。

假设区域 $R_{\alpha}, \alpha \in \{obj, bak\}$ 的灰度平均值为 μ_{α} ,区域 R_{α} 内的所有像素的灰度值服从均值为 μ_{α} ,方差为 σ_{α} 的 1 维高斯分布,即

$$P_{\alpha} = \frac{1}{2\pi\sigma_{\alpha}} \exp\left(-\frac{(R_{\alpha} - \mu_{\alpha})^2}{2\sigma_{\alpha}^2}\right), \alpha \in \{obj, bak\} \quad (18)$$

4 水平集实现

水平集方法是求解几何曲线演化的一种新颖方法,它是以隐含的方式来表达平面闭合曲线,避免了对闭合曲线演化过程的跟踪,将曲线演化转化成求解数值偏微分方程的问题。而且水平集方法避免了演化曲线的参数化过程,因此,它能够很好地处理界面在运动中的拓扑形变。这种方法自 Osher^[9]提出后,得到了广泛的应用。基本思想如下:

若运动界面为 Γ ,其运动速度为 F 。水平集方法就是将运动界面 Γ 嵌入到比运动界面高 1 维的光滑函数 $\varphi(X)$ 中,即 $\Gamma = \{X | \varphi(X, t) = 0\}$ 。通过不断地更新函数 $\varphi(X)$,从而隐含地对运动界面 Γ

进行演化。其演化过程满足以下基本偏微分方程：

$$\frac{\partial \varphi}{\partial t} = F |\nabla \varphi| \quad (19)$$

其中, $|\nabla \varphi|$ 是水平集函数的梯度范数,界面的速度 F 包括 (1) 一种是与曲率相关的扩散项,相当于曲线的内力,起到保持曲线的光滑性 (2) 一种是为曲线演化提供动力支持的对流项,以及使速度停止在目标边缘的速度衰减因子,这相当于图像力。

因此式(16)可用水平集重新表述为

$$\frac{d\varphi}{dt} = [\log P_{\text{obj}}(I^{k+1} | I^k, C^k) - \log P_{\text{bak}}(I^{k+1} | I^k, C^k) - \mu\kappa - \nu(g(\|\nabla I^{k+1}\|)\kappa - \nabla g(\|\nabla I^{k+1}\|))] |\nabla \varphi| \quad (20)$$

其中,

$$\kappa = \nabla \frac{\nabla \varphi}{\|\nabla \varphi\|} = \frac{\varphi_{xx}\varphi_y^2 - 2\varphi_x\varphi_y\varphi_{xy} + \varphi_{yy}\varphi_x^2}{(\varphi_x^2 + \varphi_y^2)^{3/2}}$$

本文数值实现采用有限差分法。根据文献[9]具体为

$$\varphi_{i,j}^{k+1} = \varphi_{i,j}^k + \Delta t [\max(F, \rho) \nabla^+ + \min(F, \rho) \nabla^- - \mu\kappa_{i,j}^k ((D_{i,j}^{ox})^2 + (D_{i,j}^{oy})^2)^{1/2} - \nu G] \quad (21)$$

其中,

$$F = \log P_{\text{obj}}(I^{k+1} | I^k, C^k) - \log P_{\text{bak}}(I^{k+1} | I^k, C^k) \quad (22)$$

$$G = g(\|\nabla I^{k+1}\|)\kappa - \nabla g(\|\nabla I^{k+1}\|) \quad (23)$$

∇^+ 和 ∇^- 的形式：

$$\nabla^+ = [\max(D_{i,j}^{-x}, \rho)^2 + \min(D_{i,j}^{+x}, \rho)^2 + \max(D_{i,j}^{-y}, \rho)^2 + \min(D_{i,j}^{+y}, \rho)^2]^{1/2} \quad (24)$$

$$\nabla^- = [\min(D_{i,j}^{-x}, \rho)^2 + \max(D_{i,j}^{+x}, \rho)^2 + \min(D_{i,j}^{-y}, \rho)^2 + \max(D_{i,j}^{+y}, \rho)^2]^{1/2} \quad (25)$$

式中 $D_{i,j}^{-x}, D_{i,j}^{+x}, D_{i,j}^{ox}$ 和 $D_{i,j}^{-y}, D_{i,j}^{+y}, D_{i,j}^{oy}$ 分别为水平集函数 φ 在 x 方向和 y 方向的前向,后向和中心差分符号：

$$D_{i,j}^{-x}\varphi_{i,j}^k = \varphi_{i,j}^k - \varphi_{i-1,j}^k$$

$$D_{i,j}^{+x}\varphi_{i,j}^k = \varphi_{i+1,j}^k - \varphi_{i,j}^k$$

$$D_{i,j}^{-y}\varphi_{i,j}^k = \varphi_{i,j}^k - \varphi_{i,j-1}^k$$

$$D_{i,j}^{+y}\varphi_{i,j}^k = \varphi_{i,j+1}^k - \varphi_{i,j}^k$$

$$D_{i,j}^{ox}\varphi_{i,j}^k = (\varphi_{i+1,j}^k - \varphi_{i-1,j}^k) / 2$$

$$D_{i,j}^{oy}\varphi_{i,j}^k = (\varphi_{i,j+1}^k - \varphi_{i,j-1}^k) / 2$$

5 实验

使用 Pentium 900、256M 内存的 Window2000 操作系统,对一段视频中的运动目标进行了跟踪实验。

序列图像的大小是 320×240 。

算法的主要步骤如下：

(1) 在第 1 帧图像 I^1 中,初始化 φ^0 ,所对应的初始轮廓为 C^0 ;

(2) 计算出有初始曲线所分成的区域 R_{in} 和 R_{out} 的平均灰度值 μ_{obj} 和 μ_{bak} ,并由此计算出条件概率密度 $P_{\text{obj}}(\varphi^0) P_{\text{bak}}(\varphi^0)$;

(3) 计算出相应的梯度图像 $\|\nabla I^k\|$;

(4) 根据式(22),通过水平集求解偏微分方程,从而得到运动目标在图像 I^1 的轮廓 C^1 ;

(5) 在图像中 $I^k, k=2, 3, \dots$,将上一帧图像中运动目标的轮廓 C^{k-1} ,作为本帧的初始轮廓,继续执行第 2 步,直至所要跟踪的最后一帧图像;

参数设置为

$$\sigma_g = \sigma_{\text{obj}} = \sigma_{\text{bak}} = 1 \quad \mu = 0.02 \times 225^2 \quad \Delta t = 0.1$$

图 1 是从图像序列中抽取的第 1、5、10、20、30、40 帧的原始图和跟踪结果,平均每帧计算耗时 5.62spf。由图 1 可见,本文提出的方法可以很好地

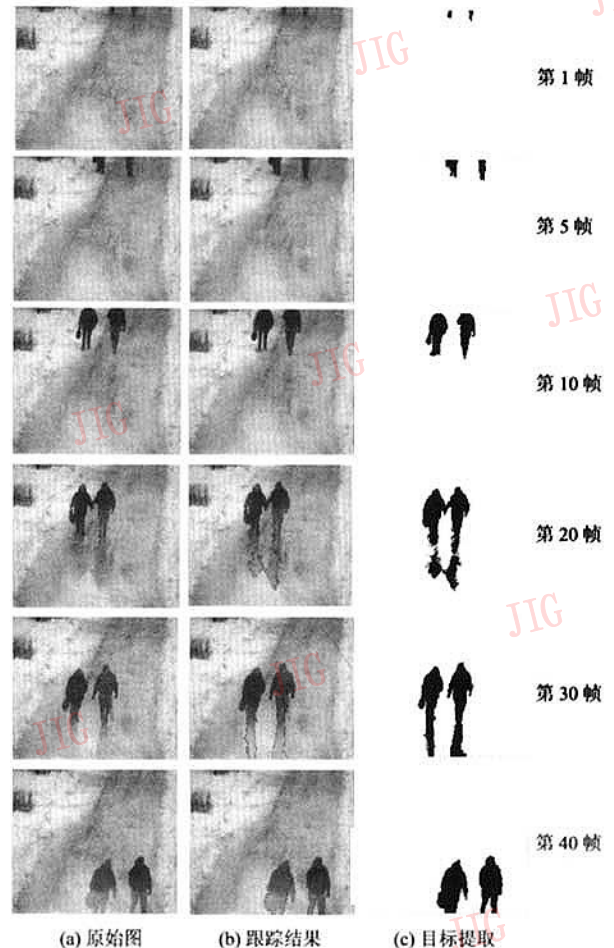


图 1 图像序列的跟踪结果

Fig. 1 Tracked result in images sequence

逼近运动目标的轮廓,并且可以自然地处理多目标问题。此外,这种算法在解偏微分方程时,只需2~3次迭代即可,效率比较高。同时,从第20帧和30帧的结果中也可以看出,这种方法将目标的影子也检测出来了,这是由于该区域的灰度值发生变化而引起的。可以采取一些措施来区分目标和阴影,比如目标和阴影的几何关系等。关于如何消除阴影,现在是一个热点研究内容。

6 结 论

本文首先根据贝叶斯理论,建立了图像序列中运动目标的跟踪模型,然后对模型分析,巧妙与基于区域的活动轮廓模型建立对应的关系,将问题的求解转化为能量的最小化问题,并通过水平集方法进行数值实现。

实验结果表明,这种方法不仅可以对多个运动目标跟踪且非常好地逼近运动目标的轮廓,而且能够自然地处理运动目标的拓扑形变。然而,其也存在一些缺陷,当背景比较复杂时,会检测到虚假目标。实验中参数的选择与所要处理的图像有很大的关联,如何根据图像特征使参数自适应地变化。此外,算法效率有所提高,但仍很难到达实时处理要求,这些都是今后研究所要考虑的。

参考文献 (References)

1 Kass M, Witkin A, Terzopoulos D. Snakes: active contour models

[J]. International Journal of Computer Vision, 1987, 1(4):321~331.

2 Caselles V, Kimmel R, Sapiro G. Geodesic active contours[J]. International Journal of Computer Vision, 1997, 22(1):61~79.

3 Zhu S C, Yuille A L. Region competition: unifying snake/balloon, region growing and Bayes/MDL/energy for multi-band image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(9):884~900.

4 Jehan-Besson S, Barlaud M, Aubert G. Video object segmentation using Eulerian region-based active contours [A]. In: IEEE International Conference in Computer Vision [C], Vancouver, Canada. 2001, 1:353~360.

5 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.

6 Chan T F, Vese L A. Active contours without edges[J]. IEEE Transactions on Image Processing, 2001, 10(2):266~277.

7 Mumford D, Shah J. Optimal approximations by piecewise smooth functions and associated variational problems[J]. Communications on Pure Applied Mathematics, 1989, 42(4):577~685.

8 Mansouri A. Region tracking via level set PDEs without motion computation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7):947~961.

9 Osher S J, Sethian J A. Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations [J]. Journal of Computational Physics, 1988, 79(1):12~49.