

中图法分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2011)02-0258-09

论文索引信息: 李培华, 肖莉娟. 基于 Mean Shift 的相似性变换和仿射变换目标跟踪算法[J]. 中国图象图形学报, 2011, 16(2): 258-266

基于 Mean Shift 的相似性变换和仿射变换目标跟踪算法

李培华, 肖莉娟

(黑龙江大学计算机科学技术学院, 哈尔滨 150080)

摘要: 传统的 Mean Shift (MS) 算法只能对发生平移和尺度变化的目标进行跟踪, 而对于具有相似性变换或者更复杂的仿射变换的目标跟踪效果很不理想或无法跟踪。为了解决这一问题, 提出了两种基于 MS 的改进算法。第一种算法针对仿射变换, 根据奇异值分解理论, 仿射变换矩阵可以分解成两个旋转矩阵和一个对角矩阵的乘积, 在此基础上建模了一种新的候选目标模型。通过 Bhattacharyya 系数将目标跟踪问题转化成以仿射变换参数为变量的最优化问题, 推导相关参量的一阶偏导数并令其为零从而得出相对于仿射变换的 MS 算法。另外, 针对进行相似性变换的目标也提出了一种新的候选目标模型, 并用类似的梯度下降算法估计目标的平移向量和旋转角度。实验结果表明, 提出的算法能够跟踪具有相似性变换或仿射变换的目标, 比传统的 MS 算法具有更好的跟踪性能。

关键词: 目标跟踪; Mean Shift 算法; 仿射变换; 相似性变换

Mean Shift based object tracking with similarity and affine transformations

Li Peihua, Xiao Lijuan

(College of Computer Science and Technology, Heilongjiang University, Harbin 150080 China)

Abstract: Traditional Mean Shift (MS) algorithm can only follow objects with translation and scale change, and fails to handle objects with similarity transformation or complex affine transformation. To address this problem, the paper presents two improved algorithms. The first one focuses on the affine motion. According to the theory of Singular Value Decomposition, the affine matrix can be factored into product of two rotation matrixes and one diagonal matrix, based on which a new candidate model is proposed. With Bhattacharyya coefficient as a similarity function, the object tracking is formulated as an optimization problem, and the corresponding MS algorithm can be derived by calculating the first derivative of the similarity function with respect to affine parameters and setting them to be zero. Furthermore, a new candidate model is proposed that handles similarity transformation, and the corresponding MS algorithm can be obtained that estimates the translation vector and rotation angle. Experimental results show that, the proposed algorithms can track objects with similarity or affine transformations, and have better tracking performance than the traditional one.

Keywords: Object tracking; Mean Shift algorithm; affine transformation; similarity transformation

0 引言

基于颜色信息的目标跟踪方法在视频监控、智

能人机交互、图像和视频编码以及智能机器人等许多研究领域中具有广泛地应用。在众多的颜色跟踪方法中, 基于 Mean Shift (MS) 的目标跟踪算法由于具有实时性、鲁棒性和易于实现等优点, 近年来吸引

收稿日期: 2009-04-01; 修回日期: 2009-10-16

基金项目: 国家自然科学基金项目(60673110, 60973080); 教育部科学技术研究重点项目(210063); 黑龙江省新世纪优秀人才计划项目(1153-NCET-002); 黑龙江大学高层次人才(创新团队)支持计划(Hdtd 2010-07)。

第一作者简介: 李培华(1971—), 男, 教授, 博士。主要研究方向是图像处理 and 模式识别。E-mail: peihualijia@sina.com。

了越来越多学者的研究兴趣^[1-13]。

Comaniciu 等人将 MS 算法成功地应用于目标跟踪领域^[1-2],他们提出用非参量的概率密度建模目标的颜色分布,用目标模型和候选目标模型之间的 Bhattacharyya 系数作为代价函数来度量概率密度的相似程度。通过令代价函数的梯度向量为零从而得到 MS 向量,MS 向量一般能够在很少的几步迭代后收敛到代价函数的极值点,即目标的真实位置。Comaniciu 等人的这一开创性的工作吸引了广泛的关注。在这一算法的基础上,众多新颖的改进算法不断涌现^[3-13]。这些算法或者致力于改善跟踪性能(跟踪准确性或鲁棒性)^[3,6-7],或者着重解决自适应核宽变化问题^[10]、多目标跟踪或关节目标跟踪问题^[8-9],或者研究具有复杂运动或形状的目标跟踪问题^[12-13]。

传统 MS 算法是一种梯度下降算法,容易收敛到代价函数的局部极值点,为此 Shen 等提出了一种退火 MS 算法^[3],顺序地减少带宽并进行 MS 迭代从而寻找代价函数的全局极值点。Zhao 等人提出用 Earth Mover's Distance(EMD)度量目标模型和候选模型之间的相似程度并推导出相应的 MS 算法^[4],以解决 Bhattacharyya 系数对由于光照变化引起的颜色模型变化敏感的问题。Han 等人提出一种序列核密度近似的方法^[5],在线地估计并更新目标模型中的高斯核的均值向量和协方差矩阵。李培华提出了一种改进 MS 算法提高目标跟踪性能^[6];同时还提出了用聚类的方法自适应地剖分目标颜色空间,建立新的目标模型^[7]。

Park 等人将 MS 算法和置信传播结合起来^[8],用来进行多目标跟踪和关节人体跟踪。Fan 等人通过子空间约束的方式使多个 MS 跟踪器互相协作^[9],从而能够跟踪关节目标或变形目标。Collins 等人将 Lindeberg 尺度的空间理论和 MS 算法相结合^[10],以解决传统的 MS 算法不能自适应变化核函数带宽的缺陷。在 Elgammal 等人提出的联合的特征-空间概率密度的工作基础上^[11],Zhang 等人将旋转信息引入到模型中并用 L_2 范数来度量概率函数的相似性^[12],该算法能够跟踪旋转目标,但是跟踪速度很慢,无法适应实时跟踪需要。Yilmaz 引入了一种改进的 Level Set 函数作为一种非对称的核函数^[13],该方法能够跟踪具有复杂形状的目标。

当目标进行复杂的运动从而发生显著的姿态变化,将目标建模为直立椭圆的 MS 算法^[1-7,9-11],跟踪

效果很不理想甚至跟踪失败。本文的工作建立在 Comaniciu 等人的工作^[1-2]基础上,旨在解决具有复杂运动的目标实时跟踪问题。将目标的仿射运动模型引入到核函数中,将目标的跟踪问题转化为仿射参量优化问题,通过代价函数的一阶偏导数为零推导出相应的 MS 算法,该算法能够跟踪具有仿射变化的目标。另外,对于能够用相似性变换描述的目标运动,即当目标仅具有平移、旋转运动和 x, y 方向相同的尺度变化时,我们用类似的方法得到了相应的 MS 跟踪算法。

1 基于 MS 算法的仿射运动目标跟踪

假设目标区域用宽和高分别为 $2h_x$ 和 $2h_y$,中心位于坐标原点的矩形表示,该目标区域像素总数为 n ,将颜色空间分成 m 个直方图区间,则目标模型可以表示为

$$\mathbf{p} = [p_1, p_2, \dots, p_m]^T \quad (1)$$

式中任意分量 p_u 定义如下:

$$p_u = C \sum_{i=1}^n k \left(\frac{x_i^2}{h_x^2} + \frac{y_i^2}{h_y^2} \right) \delta_{ui} \quad u = 1, 2, \dots, m$$

式中, C 是满足 $\sum_{u=1}^m p_u = 1$ 的归一化常数;如果点 (x_i, y_i) 的像素颜色向量属于第 u 个直方图区间,克罗内克函数 $\delta_{ui} = 1$, 否则 $\delta_{ui} = 0$; $k(\cdot)$ 是核函数。与文献[1]相同,这里采用 Epanechnikov 核函数,此时核函数的导函数是常值函数,即 $g(\cdot) = k'(\cdot) = \text{常数}$ 。

对于进行一般仿射变换的目标,其变换前后目标的位置坐标可以描述成如下矩阵形式:

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & x \\ a_{21} & a_{22} & y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

式中 (x'_i, y'_i) 是仿射变换后的坐标, $\mathbf{z} = (x, y)$ 是平移向量,上式用分块矩阵写成:

$$\mathbf{X}' = \mathbf{H}_A \mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{Z}^T \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{X} \quad (2)$$

由矩阵奇异值分解理论^[14]知:正定矩阵 \mathbf{A} 可以分解为

$$\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{V}^T \quad (3)$$

式中 $\mathbf{D} = \text{diag}\{1/\lambda_1, 1/\lambda_2\}$ 为对角阵, $1/\lambda_1, 1/\lambda_2$ 为矩阵 \mathbf{A} 的正奇异值, \mathbf{U} 为标准正交矩阵,可写成旋转矩阵形式^[15]

$$U = R(\phi) = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}$$

标准正交矩阵 $V = R(\theta)$ 具有类似的形式。

由式(2)(3), 并注意到旋转矩阵 $R(\phi)$ 和 $R(\theta)$ 的正交性, 可得:

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x'_i - x \\ y'_i - y \end{bmatrix}$$

由上式得到:

$$\begin{cases} x_i = a(x'_i - x) + b(y'_i - y) \\ y_i = c(x'_i - x) + d(y'_i - y) \end{cases}$$

式中,

$$\begin{cases} a = \lambda_1 \cos \theta \cos \phi - \lambda_2 \sin \theta \sin \phi \\ b = \lambda_1 \sin \theta \cos \phi + \lambda_2 \cos \theta \sin \phi \\ c = -\lambda_1 \cos \theta \sin \phi - \lambda_2 \sin \theta \cos \phi \\ d = -\lambda_1 \sin \theta \sin \phi + \lambda_2 \cos \theta \cos \phi \end{cases}$$

因此候选目标模型可以表示为

$$q(z, \theta, \phi, \lambda_1, \lambda_2) = [q_1(z, \theta, \phi, \lambda_1, \lambda_2), \dots, q_m(z, \theta, \phi, \lambda_1, \lambda_2)]^T \quad (4)$$

式中任意分量 q_u 具有如下形式:

$$q_u(z, \theta, \phi, \lambda_1, \lambda_2) = C_h \sum_{i=1}^{n_h} h \left(\frac{(a(x'_i - x) + b(y'_i - y))^2}{h_x^2} + \frac{(c(x'_i - x) + d(y'_i - y))^2}{h_y^2} \right) \delta_{ui}$$

模型之间的相似性由下式来度量:

$$S(p, q(\cdot)) = \sum_{u=1}^m \sqrt{p_u q_u(z, \theta, \phi, \lambda_1, \lambda_2)} \quad (5)$$

跟踪一般的仿射变换目标, 实际上就是估计平移向量 $z = (x, y)$, 两个旋转角度 θ, ϕ , 和两个尺度参数 λ_1, λ_2 , 使得式(5)中代价函数 $S(p, q(\cdot))$ 达到最大。

1.1 平移向量的估计

在式(5)中分别对 x 和 y 求偏导数可得

$$\begin{aligned} \frac{\partial S(p, q(\cdot))}{\partial x} &\propto \sum_{i=1}^{n_h} \omega_i \left\{ \left(\frac{a^2}{h_x^2} + \frac{c^2}{h_y^2} \right) (x'_i - x) + \left(\frac{ab}{h_x^2} + \frac{cd}{h_y^2} \right) (y'_i - y) \right\} \\ \frac{\partial S(p, q(\cdot))}{\partial y} &\propto \sum_{i=1}^{n_h} \omega_i \left\{ \left(\frac{b^2}{h_x^2} + \frac{d^2}{h_y^2} \right) (y'_i - y) + \left(\frac{ab}{h_x^2} + \frac{cd}{h_y^2} \right) (x'_i - x) \right\} \end{aligned}$$

式中权重 $\omega_i = \sum_{u=1}^m \sqrt{p_u / q_u(\cdot)} \delta_{ui}$ 。分别令

$$\frac{\partial S(p, q(\cdot))}{\partial x} = 0$$

$$\frac{\partial S(p, q(\cdot))}{\partial y} = 0$$

可得

$$\begin{aligned} x &= \frac{\sum_{i=1}^{n_h} x_i \omega_i g(\cdot)}{\sum_{i=1}^{n_h} \omega_i g(\cdot)} \\ y &= \frac{\sum_{i=1}^{n_h} y_i \omega_i g(\cdot)}{\sum_{i=1}^{n_h} \omega_i g(\cdot)} \end{aligned} \quad (6)$$

1.2 两个旋转角度的估计

在式(5)中分别对旋转角度 θ, ϕ 求偏导数得

$$\begin{aligned} \frac{\partial S(p, q(\cdot))}{\partial \theta} &\propto \sum_{i=1}^{n_h} \omega_i \left\{ - \left(\frac{a_1 a_2}{h_x^2} + \frac{a_3 a_4}{h_y^2} \right) \cos 2\theta + \frac{1}{2} \left(\frac{a_1^2 - a_2^2}{h_x^2} + \frac{a_3^2 - a_4^2}{h_y^2} \right) \sin 2\theta \right\} \\ \frac{\partial S(p, q(\cdot))}{\partial \phi} &\propto \sum_{i=1}^{n_h} \omega_i \left\{ - \left(\frac{b_1 b_2}{h_x^2} + \frac{b_3 b_4}{h_y^2} \right) \cos 2\phi + \frac{1}{2} \left(\frac{b_1^2 - b_2^2}{h_x^2} + \frac{b_3^2 - b_4^2}{h_y^2} \right) \sin 2\phi \right\} \end{aligned}$$

式中,

$$\begin{cases} a_1 = \lambda_1 (x'_i - x) \cos \phi + \lambda_2 (y'_i - y) \sin \phi \\ a_2 = -\lambda_2 (x'_i - x) \sin \phi + \lambda_1 (y'_i - y) \cos \phi \\ a_3 = -\lambda_1 (x'_i - x) \sin \phi + \lambda_2 (y'_i - y) \cos \phi \\ a_4 = -\lambda_2 (x'_i - x) \cos \phi - \lambda_1 (y'_i - y) \sin \phi \\ b_1 = \lambda_1 (x'_i - x) \cos \theta + \lambda_1 (y'_i - y) \sin \theta \\ b_2 = -\lambda_2 (x'_i - x) \sin \theta + \lambda_2 (y'_i - y) \cos \theta \\ b_3 = -\lambda_2 (x'_i - x) \sin \theta + \lambda_2 (y'_i - y) \cos \theta \\ b_4 = -\lambda_1 (x'_i - x) \cos \theta - \lambda_1 (y'_i - y) \sin \theta \end{cases}$$

令 $\frac{\partial S(p, q(\cdot))}{\partial \theta} = 0$, 得到

$$\theta = \arctan \frac{-c_2 \pm \sqrt{c_2^2 + 4c_1^2}}{2c_1} \quad (7)$$

式中,

$$c_1 = \frac{\lambda_1 \lambda_2}{2} \left(\frac{1}{h_y^2} - \frac{1}{h_x^2} \right) \sin(2\phi) \sum_{i=1}^{n_h} \omega_i [(x'_i - x)^2 - (y'_i - y)^2] + \left[\left(\frac{\lambda_1^2}{h_x^2} - \frac{\lambda_2^2}{h_y^2} \right) \cos^2 \phi + \left(\frac{\lambda_1^2}{h_y^2} - \frac{\lambda_2^2}{h_x^2} \right) \sin^2 \phi \right] \times$$

$$\sum_{i=1}^{n_h} \omega_i (x'_i - x) (y'_i - y)$$

$$c_2 = 2\lambda_1\lambda_2 \left(\frac{1}{h_x^2} - \frac{1}{h_y^2} \right) \sin(2\phi) \sum_{i=1}^{n_h} \omega_i (x'_i - x)(y'_i - y) +$$

$$\left[\left(\frac{\lambda_1^2}{h_x^2} - \frac{\lambda_2^2}{h_y^2} \right) \cos^2\phi + \left(\frac{\lambda_1^2}{h_x^2} - \frac{\lambda_2^2}{h_y^2} \right) \sin^2\phi \right] \times$$

$$\sum_{i=1}^{n_h} \omega_i [(x'_i - x)^2 - (y'_i - y)^2]$$

令 $\frac{\partial S(\mathbf{p}, \mathbf{q}(\cdot))}{\partial \phi} = 0$, 得到

$$\phi = \arctan \frac{-d_2 \pm \sqrt{d_2^2 + 4d_1^2}}{2d_1} \quad (8)$$

式中

$$d_1 = \frac{\lambda_1\lambda_2}{2} \left(\frac{1}{h_y^2} - \frac{1}{h_x^2} \right) \sin(2\theta) \sum_{i=1}^{n_h} \omega_i [(x'_i - x)^2 - (y'_i - y)^2] +$$

$$\lambda_1\lambda_2 \left(\frac{1}{h_x^2} - \frac{1}{h_y^2} \right) \sin(2\theta) \sum_{i=1}^{n_h} \omega_i (x'_i - x)(y'_i - y) \times$$

$$\sum_{i=1}^{n_h} \omega_i (x'_i - x)(y'_i - y)$$

$$d_2 = \left(\frac{1}{h_x^2} - \frac{1}{h_y^2} \right) \left[(\lambda_1^2 \cos^2\theta - \lambda_2^2 \sin^2\theta) \sum_{i=1}^{n_h} \omega_i (x'_i - x)^2 + \right.$$

$$\left. (\lambda_1^2 \sin^2\theta - \lambda_2^2 \cos^2\theta) \sum_{i=1}^{n_h} \omega_i (y'_i - y)^2 + \right.$$

$$\left. (\lambda_1^2 + \lambda_2^2) \sin(2\theta) \sum_{i=1}^{n_h} \omega_i [(x'_i - x)^2 - (y'_i - y)^2] \right]$$

两个旋转角度 θ 和 ϕ 分别有两个解, 如果目标的宽度大于高度, 则在式(7)(8)中, 取“+”, 否则取“-”。

1.3 跟踪算法

在仿射变换中, 两个尺度无法用优化的方式求出, 因此本文采用与文献[1]类似的方法: 采用离散尺度因子的方法, 即分别搜索 $\lambda_1, \lambda_1 \pm 0.05\lambda_1$ 和 $\lambda_2, \lambda_2 \pm 0.05\lambda_2$, 最后选择相似性最大的尺度。跟踪算法(算法1)如下:

对目标的状态进行初始化, 计算目标模型 \mathbf{p}

while 获得新的一帧图像 do

for 对 λ_1, λ_2 的每一对离散尺度 do

repeat

根据式(4)计算候选目标模型 $\mathbf{q}(\cdot)$;

根据式(6)更新平移向量 \mathbf{z} , 计算平移向量变化的 L_2 范数 $\Delta\mathbf{z}$;

根据式(7)更新 θ , 计算变化量的绝对值 $\Delta\theta$;

根据式(8)更新 ϕ , 计算 ϕ 变化量绝对值 $\Delta\phi$;

until

$\Delta\mathbf{z} \leq \varepsilon_z, \Delta\theta \leq \varepsilon_\theta, \Delta\phi \leq \varepsilon_\phi$ 或迭代次数 > 10

end for

选择具有最大相似性的尺度参量

end while

2 基于 MS 跟踪具有相似性变换的目标

对于具有相似性变换的目标, 其坐标变换可描述成下述矩阵形式:

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda/\cos\varphi & -\lambda/\sin\varphi & x \\ \lambda/\sin\varphi & \lambda/\cos\varphi & y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

式中 (x'_i, y'_i) 是相似性变换后目标的坐标, $\mathbf{z} = (x, y)$ 是平移分量, $1/\lambda$ 是尺度参量, φ 是旋转参量。由上式可以得到发生相似性变换后候选模型为

$$\mathbf{q}(\mathbf{z}, \varphi, \lambda) = [q_1(\mathbf{z}, \varphi, \lambda), \dots, q_m(\mathbf{z}, \varphi, \lambda)]^T \quad (9)$$

式中,

$$q_u(\mathbf{z}, \varphi, \lambda) = C_h \sum_{i=1}^{n_h} k \left(\frac{\lambda^2 ((x'_i - x) \cos\varphi + (y'_i - y) \sin\varphi)^2}{h_x^2} + \right.$$

$$\left. \frac{\lambda^2 (-(x'_i - x) \sin\varphi + (y'_i - y) \cos\varphi)^2}{h_y^2} \right) \delta_{ui}$$

此时目标模型和候选模型之间的相似性度量为

$$S(\mathbf{p}, \mathbf{q}(\mathbf{z}, \varphi, \lambda)) = \sum_{u=1}^m \sqrt{p_u q_u(\mathbf{z}, \varphi, \lambda)} \quad (10)$$

2.1 平移向量的估计

在式(10)中分别对 x 和 y 求偏导数可得

$$\frac{\partial S(\mathbf{p}, \mathbf{q}(\cdot))}{\partial x} \propto$$

$$\sum_{i=0}^{n_h} \omega_i \left\{ (x'_i - x) \left(\frac{\cos^2\varphi}{h_x^2} + \frac{\sin^2\varphi}{h_y^2} \right) + \right.$$

$$\left. \frac{1}{2} (y'_i - y) \left(\frac{1}{h_x^2} - \frac{1}{h_y^2} \right) \sin(2\varphi) \right\}$$

$$\frac{\partial S(\mathbf{p}, \mathbf{q}(\cdot))}{\partial y} \propto$$

$$\sum_{i=0}^{n_h} \omega_i \left\{ (y'_i - y) \left(\frac{\cos^2\varphi}{h_x^2} + \frac{\sin^2\varphi}{h_y^2} \right) + \right.$$

$$\left. \frac{1}{2} (x'_i - x) \left(\frac{1}{h_x^2} - \frac{1}{h_y^2} \right) \sin(2\varphi) \right\}$$

令 $\partial S(\mathbf{p}, \mathbf{q}(\cdot)) / \partial x = 0, \partial S(\mathbf{p}, \mathbf{q}(\cdot)) / \partial y = 0$ 得

$$x = \frac{\sum_{i=1}^{n_h} x_i \omega_i g(\cdot)}{\sum_{i=1}^{n_h} \omega_i g(\cdot)} \quad (11)$$

$$y = \frac{\sum_{i=1}^{n_h} y_i \omega_i g(\cdot)}{\sum_{i=1}^{n_h} \omega_i g(\cdot)}$$

2.2 旋转角度的估计

在式(10)中对 φ 求偏导数可得

$$\frac{\partial S(\mathbf{p}, \mathbf{q}(\cdot))}{\partial \varphi} \propto \sum_{i=1}^{n_h} \left\{ \omega_i \left(\frac{1}{h_x^2} - \frac{1}{h_y^2} \right) \times \right.$$

$$\left. \left[((x_i - x)^2 - (y_i - y)^2) \sin(2\varphi) - \right. \right.$$

$$\left. \left. 2(x_i - x)(y_i - y) \cos(2\varphi) \right] \right\}$$

令 $\partial S(\mathbf{p}, \mathbf{q}(\cdot)) / \partial \varphi = 0$, 可得

$$\varphi = \arctan \frac{\sum_{i=1}^{n_h} \omega_i [(y_i - y)^2 - (x_i - x)^2] \pm \sqrt{\Delta}}{2 \sum_{i=1}^{n_h} \omega_i (x_i - x)(y_i - y)} \quad (12)$$

式中判别式 Δ 具有如下形式

$$\Delta = \left(\sum_{i=1}^{n_h} \omega_i ((x_i - x)^2 - (y_i - y)^2) \right)^2 +$$

$$4 \left(\sum_{i=1}^{n_h} \omega_i (x_i - x)(y_i - y) \right)^2$$

在式(12)中, 旋转角度 φ 的取法与第 1.2 节中角度 θ 和 ϕ 的取法类似。

2.3 跟踪算法

在相似性变换中, 尺度参量无法用优化的方式求出, 这里同样采用与文献[1]类似的方法: 采用离散尺度的方法, 即分别搜索 $\lambda, \lambda \pm 0.05\lambda$, 选择相似性最大的尺度。能够跟踪具有相似性变换的目标的跟踪算法(算法 2)如下:

```

对目标的状态进行初始化, 计算目标模型  $\mathbf{p}$ 
while 获得新的一帧图像 do
  for 对  $\lambda$  的每一个离散尺度 do
    repeat
      根据式(9)计算候选目标模型  $\mathbf{q}(\cdot)$ ;
      根据式(11)更新平移向量  $\mathbf{z}$ , 计算平移向量变化的  $L_2$  范数  $\Delta \mathbf{z}$ ;
      根据式(12)更新  $\varphi$ , 计算变化量的绝对值  $\Delta \varphi$ ;
    until
  
```

$\Delta \mathbf{z} \leq \varepsilon_z, \Delta \varphi \leq \varepsilon_\varphi$ 或迭代次数 > 10

end for

选择具有最大相似性的尺度参量

end while

3 计算复杂性分析

在传统的 MS 跟踪算法中, 令 N_λ 是为了适应目标尺寸变化而搜索的尺度数目, N_r 为梯度下降算法的平均迭代次数, c_h 为加权直方图的计算代价, c_t 为计算平移向量的代价, c_s 为计算 Bhattacharyya 系数的代价, 则传统 MS 跟踪算法的总计算代价为

$$C_{MS} = N_\lambda (N_r (c_h + c_t) + c_s) \quad (13)$$

在基于 MS 的仿射运动目标跟踪算法(算法 1)中, 对于 λ_1 和 λ_2 的每一个离散尺度组合, 在每一步迭代过程中, 除了需要计算平移向量外, 还需要分别计算两个角度 θ 和 ϕ 。令 N_{λ_1} 和 N_{λ_2} 分别是为了适应离散尺度 λ_1 和 λ_2 变化而搜索的尺度数目, c_θ 为根据式(7)计算角度 θ 的代价, c_ϕ 为根据式(8)计算角度 ϕ 的代价, 则算法 1 的总计算代价为

$$C_{AF} = N_{\lambda_1} N_{\lambda_2} (N_r (c_h + c_t + c_\theta + c_\phi) + c_s) \quad (14)$$

与传统的 MS 跟踪算法相比, 具有相似性变换的目标跟踪算法(算法 2)在每一步迭代过程中, 除了更新平移向量外, 还需要更新旋转角度 φ 。令 c_φ 为根据式(12)计算角度 φ 的代价, 则算法 1 的总计算代价为

$$C_{SI} = N_\lambda (N_r (c_h + c_t + c_\varphi) + c_s) \quad (15)$$

在所有的算法中离散尺度的搜索数目相同(见 1.3 节和 2.3 节), 即 $N_\lambda = N_{\lambda_1} = N_{\lambda_2} = 3$ 。从式(13)(15)的对比可以看出, 算法 2 与传统的 MS 跟踪算法具有大致相同的计算量。而根据式(13)(14)的比较可知算法 1 的计算量比传统的 MS 跟踪算法的计算量有显著的增加, 尽管如此, 算法 1 仍然能够实时实现。

4 实验

程序在配置了 3.2 GHz CPU 和 1.5 GB 内存的台式机 HP Compaq dx6128 MT 上调试通过, 编程环境是 Visual C++ 8.0 (Visual Studio 2005)。实验中采用 RGB 颜色空间和 $16 \times 16 \times 16$ 的直方图, 初始化是在

第 1 帧中用手工完成的,实验中 $\varepsilon_x = 0.5$ 像素, $\varepsilon_\theta = \varepsilon_\phi = \varepsilon_\psi = 0.01$ 弧度。本文采用 Epanechnikov 核函数^[1]。

4.1 仿射运动目标跟踪结果

汽车序列 (car seq) 该人工合成序列共包括 100 帧,图像尺寸为 384×288 ,草坪上的桔黄色汽车用图像处理的方法叠加到一个真实的背景图像中,该汽车遵循如下的仿射运动:平移运动是均值为 (210, 233),标准差为 10 的随机行走(random walk);旋转角度 θ 均匀地从 0° 变化到 -30° ;尺度 λ_1 均匀地从 1.0 变化到 3.0,尺度 λ_2 均匀地从 1.0 变化到 0.5;每一帧图像中使用 Matlab 软件的 imnoise 函数加入了零均值、方差为 0.05 的高斯白噪声。

传统 MS 算法和算法 1 的典型跟踪结果分别如图 1(a)(b)所示,从左至右、从上至下分别是第 1, 25, 50 和 100 帧的跟踪结果。从图 1 可见,尽管传统 MS 算法给出了目标的大致位置,可是跟踪效果很差。相比之下,算法 1 准确地估计了目标的仿射运动。当在序列图像中分别加入椒盐噪声和斑点噪声,算法 1 也都能准确跟踪,由于篇幅所限,这里没有给出相应的典型跟踪结果。在汽车序列中传统 MS 算法和算法 1 的每帧平均跟踪时间分别为 13 ms 和 40 ms。

挥臂序列 (arm seq) 该序列是使用手持式摄像机 Sony DCR-PC1000E 拍摄的,共包括 100 帧,图像尺寸为 352×288 。在该序列中,测试者穿着长袖衬衫面向摄像机挥动手臂。由于手臂具有较大的深度变化,因此相似性变换不能很好地建模手臂的运动。注意到整个衬衫具有相同的颜色和纹理,因此在跟踪手臂时跟踪器很容易被衬衫的其他部分吸引。图 2(a)(b) 分别给出了传统 MS 算法和算法 1 的典型跟踪结果,从左至右、从上至下分别是第 1、15、55 和 65 帧的跟踪结果。传统 MS 算法仅仅跟踪几帧后就目标丢失了,错误地锁定在测试者的前胸处,而算法 1 能够在整个序列中稳定地跟踪目标。在挥臂序列中算法 1 的每帧平均跟踪时间是 18 ms。

4.2 具有相似性变换的目标跟踪结果

人脸序列 (face seq) 该序列也是使用手持式摄像机 Sony DCR-PC1000E 拍摄的,共包括 160 帧,图像尺寸为 352×288 。图 3(a)(b) 分别给出了传统 MS 算法和算法 2 典型跟踪结果,从左至右、从上

至下分别为第 60、100、120 和 150 帧的跟踪结果。从图 3 中可见,两种算法都能够跟踪目标,然而传统 MS 算法始终将目标建模为直立的椭圆,因而无法给出人脸姿态的估计,而算法 2 能够估计出人脸的旋转方向。因而算法 2 适用于人机交互、人脸识别等需要估计人脸姿态的领域。传统 MS 算法和算法 2 的每帧平均跟踪时间分别为 5 ms 和 8 ms。

4.3 量化比较

FPR (false positive rate) 和 FNR (false negative rate) 是背景建模中度量前景检测准确性的常用量化指标^[16]。为了进一步量化地比较传统 MS 算法和本文所提出的算法对目标(前景)跟踪的准确程度,我们定义

$$FPR_i = \frac{\text{第 } i \text{ 帧中背景像素被误认为目标像素的数目}}{\text{第 } i \text{ 帧中目标像素总数}}$$

$$FNR_i = \frac{\text{第 } i \text{ 帧中目标像素被误认为背景像素的数目}}{\text{第 } i \text{ 帧中背景像素总数}}$$

FPR_i 和 FNR_i 越小,表明在第 i 帧图像中跟踪结果越准确;理想情况下 $FPR_i = FNR_i = 0$,表明跟踪结果没有任何误差。平均 FPR 和 FNR 分别定义为在所有帧中的 FPR 和 FNR 平均值,即

$$\overline{FPR} = \frac{1}{N} \sum_{i=1}^N FPR_i$$

$$\overline{FNR} = \frac{1}{N} \sum_{i=1}^N FNR_i$$

式中 N 表示序列中图像帧的总数。

在实验中真值 (ground truth) 是通过手工标注的方式得到的。图 4 给出了汽车序列、胳膊序列和人脸序列实验中每一帧的 FPR 和 FNR,其中红色虚线表示新算法的结果而蓝色实线表示传统 MS 跟踪算法的结果,表 1 给出了序列中所有图像帧中的平均 FPR 和平均 FNR。

针对仿射运动的目标,从图 4(a) 及表 1 可见,在汽车序列实验中算法 1 (仿射跟踪算法) 的跟踪结果远远优于传统 MS 算法;从图 4(b) 及表 1 可见,在胳膊序列实验中由于传统 MS 算法失锁(将目标丢失),其 FNR 和 FPR 都远远高于算法 1。这两个实验验证了算法 1 能够有效地跟踪具有仿射运动的目标,而传统 MS 算法却不能。针对具有相似性变换的目标,从图 4(c) 和表 1 可知,算法 2 的跟踪性能要优于传统 MS 算法。而且从 4.2 节可知,算法 2 能够而传统 MS 算法无法给出目标旋转角度的估计。

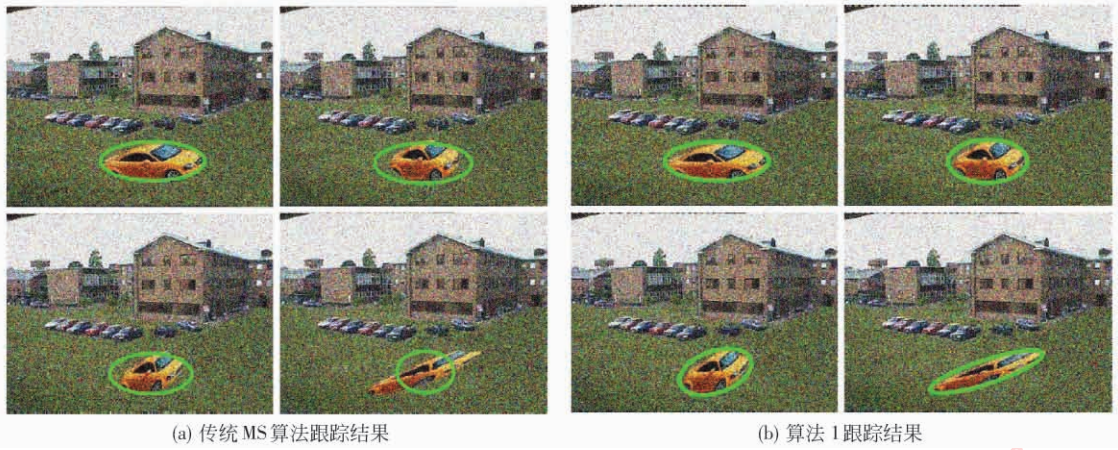


图 1 汽车序列的典型跟踪结果

Fig. 1 Some typical tracking results in car sequence



图 2 挥臂序列的典型跟踪结果

Fig. 2 Some typical tracking results in arm sequence



图 3 人脸序列的典型跟踪结果

Fig. 3 Some typical tracking results in face sequence

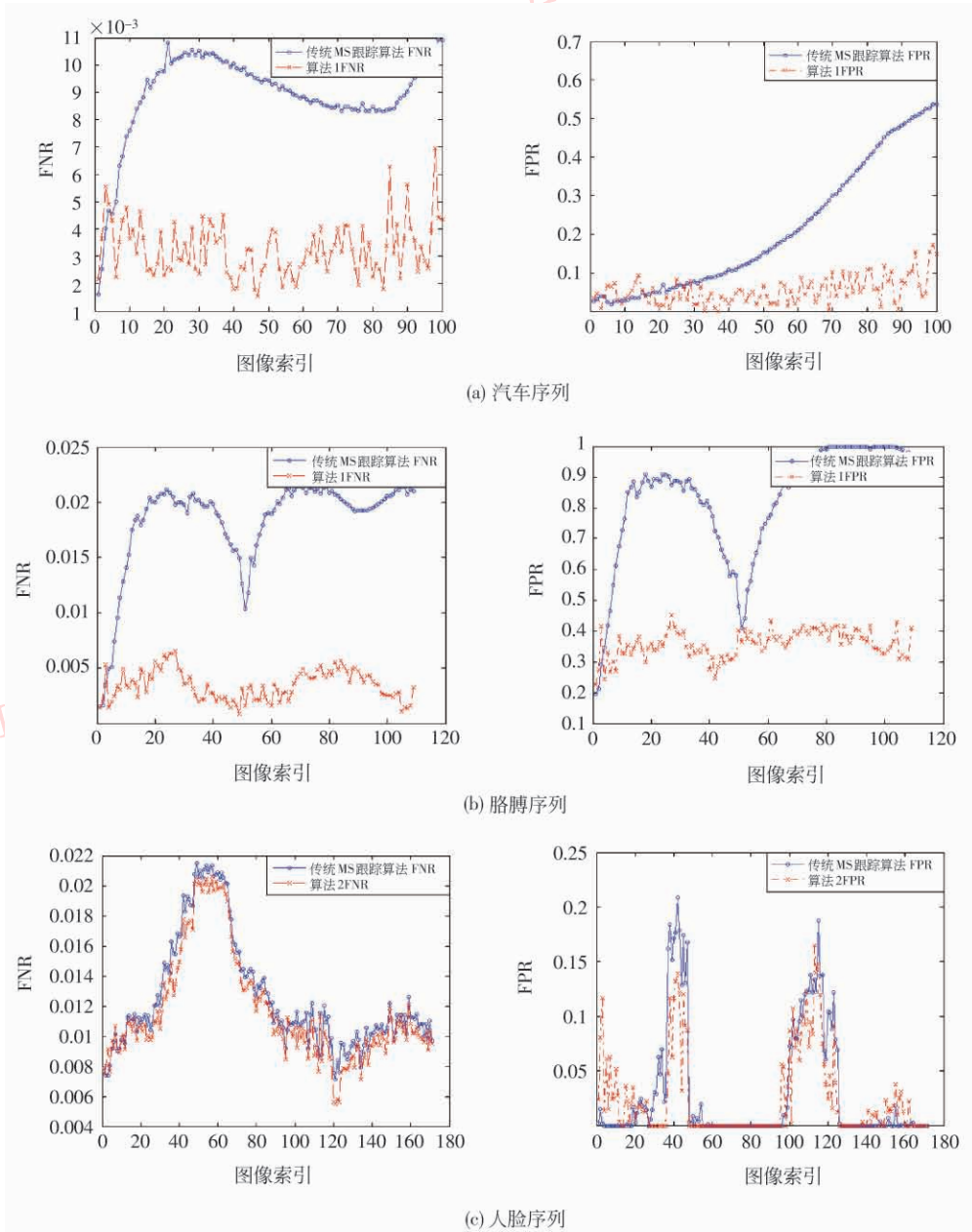


图 4 算法的量化比较

Fig. 4 Quantitative comparison between the algorithms

表 1 平均 FNR 和平均 FPR
Tab.1 Average FNR and FPR

序列	算法	FNR	FPR
汽车(仿射)	传统 MS 算法	0.008 91	0.210 1
	算法 1	0.003 25	0.053 0
胳膊(仿射)	传统 MS 算法	0.018 26	0.826 1
	算法 1	0.003 44	0.358 2
人脸(相似性)	传统 MS 算法	0.012 59	0.030 12
	算法 2	0.011 76	0.024 82

5 结 论

提出了两种基于 MS 的改进算法,分别对发生仿射变换和相似性变换的目标进行跟踪。通过将仿射变换和相似性变换引入到目标的候选模型中,将复杂运动跟踪问题转化成代价函数的优化问题。通过分别对各自的代价函数求相关参数的偏导数并令其为零,我们得到基于 MS 的仿射运动目标跟踪算

法(算法 1)和基于 MS 的相似性变换目标跟踪算法(算法 2)。在人工合成图像序列和真实图像序列中的实验表明,提出的算法能够跟踪发生复杂运动的目标,同时仍然能够实时实现。

如果近似为平面的目标在深度方向的运动与景深相比较小,即目标运动能够用相似性变换描述,则算法 2 能够很好地跟踪目标并且跟踪速度很快。算法 1 具有一般性,能够跟踪具有仿射变化的目标,尽管与传统 MS 算法相比跟踪速度较慢,但仍然能够达到实时。对于更复杂的运动,可以用射影变换描述目标的运动,用类似的方法推导出相应的 MS 算法。然而由于射影变换具有 8 个自由度^[15],因此跟踪算法会更慢因而无法达到实时,而且由于需要估计的参量过多算法可能会变得不稳定。

志谢

审稿专家的评审意见使本文理论分析更加丰富,实验内容更加充实,在此深表谢意!

参考文献 (References)

- [1] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC, USA : IEEE Press, 2000: 142-149.
- [2] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5) : 564-577.
- [3] Shen C, Brooks M J, Hengel A. Fast global kernel density mode seeking: applications to localization and tracking [J]. IEEE Transactions on Image Processing, 2007, 16(5) : 1457-1469.
- [4] Zhao Q, Brennan S, Tao H. Differential EMD tracking [C]//Proceedings of IEEE Conference on Computer Vision. Washington DC, USA : IEEE Press, 2007: 1-8.
- [5] Han B, Comaniciu D, Zhu Y, et al. Sequential kernel density approximation and its application to real-time visual tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(7) : 1186-1197.
- [6] Li Peihua. An improved Mean Shift algorithm for object tracking [J]. Acta Automatica Sinica, 2007, 33(4) : 347-354. [李培华. 一种改进的 Mean Shift 跟踪算法 [J]. 自动化学报, 2007, 33(4) : 347-354.]
- [7] Li P. An adaptive binning color model for mean shift tracking [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(9) : 1293-1299.
- [8] Park M, Liu Y, Collins R. Efficient mean shift belief propagation for vision tracking [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC, USA : IEEE Press, 2008: 1-8.
- [9] Fan Z, Yang M, Wu Y. Multiple collaborative kernel tracking [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2007, 29(7) : 1268-1273.
- [10] Collins R T. Mean-shift blob tracking through scale space [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Madison. Washington DC, USA : IEEE Press, 2003: 234-240.
- [11] Elgammal A, Duraiswami R, Davis L. Probabilistic tracking in joint feature-spatial spaces [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC, USA : IEEE Press, 2003: 781-788.
- [12] Zhang H, Huang Z, Huang W, et al. Kernel-based method for tracking objects with rotation and translation [C]//Proceedings of IEEE Conference on Pattern Recognition. Washington DC, USA : IEEE Press, 2004 : 728-731.
- [13] Yilmaz A. Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC, USA : IEEE Press, 2007: 1-6.
- [14] Zhang Xianda. Matrix Analysis and Applications [M]. Beijing: Tsinghua University Press, 2004: 344-358. [张贤达. 矩阵分析与应用 [M]. 北京: 清华大学出版社, 2004: 344-358.]
- [15] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision [M]. England: Cambridge University Press, 2000: 25-62.
- [16] Heikkilä M, Pietikäinen M. A texture based method for modeling the background and detecting moving objects [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4) : 657-662.