

基于 Hausdorff 距离的 3 维模型匹配的改进方法

陈莹 韩崇昭

(西安交通大学电信学院综合自动化研究所, 西安 710049)

摘要 通过单目灰度图像来实现已知 3 维模型移动对象的精确定位, 是基于 3 维模型的交通视觉检测与目标跟踪系统的首要环节, 也是机器视觉领域的一个重要问题。为了更好地进行图像匹配, 提出了一种带权值的 Hausdorff 距离作为 3 维模型投影和图像中物体轮廓相似性的测度, 以避免建立图像特征与模型之间的点-点对应, 这样既可减少计算量, 也可提高匹配精度。为了避免陷入局部最优, 可将一种带记忆功能的模拟退火(SA)算法引入图像模型匹配, 这样可提高匹配参数的搜索精度。实验证明, 由于 SA 算法和改进的 Hausdorff 距离相结合能有效地对 3 维模型和图像进行匹配, 从而可对具有平移、旋转的物体实现精确定位。

关键词 Hausdorff 距离 边缘提取 3 维模型匹配 模拟退火

中图分类号: TP242.6*2 TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2005)03-0326-06

Improved 3D Model Matching Based on Hausdorff Distance

CHEN Ying, HAN Chong-zhao

(Institution of Synthetic Automation, School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049)

Abstract The localization and recognition of known three-dimensional (3D) objects from single monocular intensity images is a principle component in vision-based traffic scene, and is also one of the fundamental problems in computer vision. For better image matching, the improved modified Hausdorff distance is proposed to measure the degree of similarity between 3D model and image's contour, which avoids establishing one-by-one relationship between image feature and model, greatly reduces the computational complexity, and improves the matching precision. An improved simulated annealing (SA) algorithm with a mnemonic is used for model matching. Because SA has high parallel and robust performance, it will find global optimum instead of getting into partial one. Experiments confirm that the combination of SA and improved Hausdorff distance can effectively realize matching between model and image, and localize the object that is changed in translation and rotation.

Keywords Hausdorff distance, edge detection, 3D model matching, simulated annealing

1 引言

如何通过单目灰度图像来实现已知 3 维模型的移动对象的精确定位, 是基于 3 维模型的交通视觉检测与目标跟踪的首要任务, 也是机器视觉领域的一个重要问题, 其定位的关键在于建立起目标图像与模型之间的对应关系。传统的定位方法往往需要首先从 2 维图像中提取特征, 然后与对应的 3 维模

型特征进行匹配^[1]。由于特征提取与匹配过程对遮挡、光照、噪声敏感, 且时间损耗大, 因此难以满足实际应用的需要^[2]。Hausdorff 距离可用于测量两个点集的匹配程度, 由于它无需建立模型和图像之间精确的点-点对应, 因此比其他匹配方法更具容错性^[3]。Yang Hao 等人采用一种类似 Hausdorff 距离的 PLS (point-to-line-segment) 距离实现了目标图像与模型投影的位移匹配, 由于其旋转参数是通过不在平行于 GP (ground plane) 平面的空间中的任意 3

基金项目: 国家重点基础研究发展规划资助项目 (2001CB309403)

收稿日期: 2004-04-06; 改回日期: 2004-08-09

第一作者简介: 陈莹 (1976 ~), 女, 2001 年获燕山大学机械设计与自动化专业硕士学位, 现为西安交通大学电子与信息工程学院控制科学与工程系博士研究生。研究方向为计算机视觉、目标跟踪、智能交通系统。E-mail: chen_ying_ann@sohu.com

点所组成的角的变化来获得^[4],因此该方法在图像中,首先需要搜索与 3 维模型投影中不在平行于 GP 平面的空间中的任意 3 点对应的像素点,这就要求模型初始姿态应尽量接近正确姿态,否则将大大影响匹配结果。

本文所提出的 3 维模型匹配算法是以改进的 Hausdorff 距离作为模型投影与目标图像匹配的依据。由于传统的 Hausdorff 距离依赖于一个点集和另一个点集中最不匹配点的距离,因此使得它对任何远离中心的噪声点非常敏感。本文的算法则通过简便可行的边界提取,采用带有权值的 MHD (modified Hausdorff distance) 滤除了远离边界的噪声点对模型匹配的影响;同时,为了避免在图像中寻找与模型投影的某可见点所对应的像素点,本文将模型的位移与旋转参数直接定义在 3 维空间,同时通过坐标变换与 Hausdorff 距离来建立起目标图像与 3 维模型平面投影之间的全局匹配优化函数,并通过带有记忆的模拟退火(simulated annealing, SA) 算法来寻求姿态最优解,以提高匹配的精度。

2 改进的 Hausdorff 距离

给定 2 个有限集 $M = \{m_1, m_2, \dots, m_p\}$, $T = \{t_1, t_2, \dots, t_q\}$, 则 M, T 之间的 Hausdorff 距离定义如下:

$$H(M, T) = \max(h(M, T), h(T, M)) \quad (1)$$

其中

$$\begin{aligned} h(M, T) &= \max_{m \in M} \min_{t \in T} \|m - t\|, \\ h(T, M) &= \max_{t \in T} \min_{m \in M} \|t - m\| \end{aligned} \quad (2)$$

其中, $\|\cdot\|$ 为定义在点集 M 和 T 上的某种距离范数。函数 $h(M, T)$ 称为从点集 M 到点集 T 的有向 Hausdorff 距离,若定义一个点到有限集合的距离为该点与这个集合所有点的距离的最小值,那么 $h(T, M)$ 即为点集 M 中的每个点到点集 T 的距离的最大值。由于 Hausdorff 距离 $H(M, T)$ 是取 $h(M, T)$ 和 $h(T, M)$ 的最大值,因此通过计算两个集合中的点到对方集合距离最远的点所对应的距离,即可获得 2 个点集 M 和 T 之间不匹配的程度。

由于由式(1)所定义的 Hausdorff 距离依赖于一个物体和另一个物体中最不匹配点的距离,这使得它对任何远离中心的噪声点非常敏感,为此,人们对其进行了改进。Dubuisson 和 Jain 在研究了 24 种改进形式后,指出 MHD 具有最好的性能^[5]。有向

MHD 定义为

$$h_{\text{MHD}}(M, T) = \frac{1}{P} \sum_{m_i \in M} \min_{t_j \in T} \|m_i - t_j\| \quad (3)$$

其中, P 为模型集 M 中元素的个数。

本文采用的算法对式(3)进行了改进,即利用目标图像中的主要点——轮廓点代替所有的二值像素来进行模型匹配,这样,一方面可以减少计算量,另一方面更进一步降低了噪声点对匹配的干扰^[6]。再次改进的带权 Hausdorff 距离定义为

$$\begin{aligned} \hat{h}_{\text{MHD}}(T, M) &= \frac{1}{\sum_{t_j \in T} W_{t_j}} \sum_{t_j \in T} W_{t_j} \cdot \min_{m_i \in M} \|t_j - m_i\| \\ &= \frac{1}{\sum_{t_j \in T} W_{t_j}} \sum_{t_j \in T} W_{t_j} \cdot \text{mind}(t_j, m_i) \end{aligned} \quad (4)$$

其中,权值 W_{t_j} 表示了目标图像像素点 t_j 对模型匹配的重要程度,权值的选取应遵循以下规则:

(1) 由于车辆的最外层轮廓(对应 x, y 坐标的最大和最小值)对模型匹配的影响最大,因此当 t_j 为外轮廓点时,则 W_{t_j} 取值较大,根据多次实验, W_{t_j} 通常取 15 ~ 20;

(2) 由于车身变化而引起的内轮廓点 t_j 所对应的权值 $W_{t_j} = 1$ 。

3 目标轮廓点的提取

若要利用式(4)进行模型匹配,则首先需要获取目标的轮廓点。尽管目标边界的提取方法很多^[7],但通常都比较耗时,无法适应跟踪系统实时的要求。由于本文所研究的目标是车辆,其轮廓线通常由两组平行线段组成(如图 1 所示),因此对图像进行了预处理,并用“Canny”算子提取边界后,采用如下方法提取目标的轮廓点:

(1) 从上至下,从左至右扫描 ROI (region of interest) 图像,并分别在集合 X, Y 中记下所有像素点为 1 的 x, y 坐标及其该坐标出现的次数;

(2) 分别将 X, Y 中的元素,按出现次数的多少从大到小排列,并取其前 6 位,放至集 \hat{X}, \hat{Y} ;



图 1 目标轮廓点

Fig. 1 Target's outline

(3) 再次从上至下, 从左至右扫描 ROI 图像, 若某点的 8-邻域点的坐标值满足 $x \in \hat{X}$ 或 $y \in \hat{Y}$, 则认为该点为目标轮廓点。

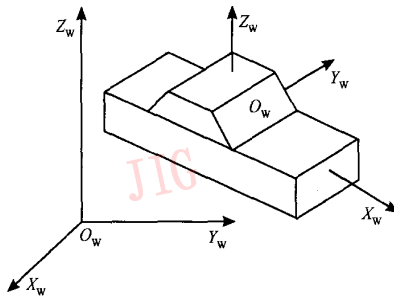
为验证上述方法的可行性, 对多幅各种姿态的目标图像进行了轮廓提取试验, 提取结果如图 1 所示, 其中矩形框表示 ROI 区域。

4 匹配优化模型

由于在 GPC (ground plane constraint) 条件下^[2,4,8], 车辆的运动被限制在地平面之上, 其自由度从 6 降到 3 (即车辆在地平面的位移 (a, b) 以及绕垂直 GP 平面的轴旋转 φ 角), 因此在本文的模型匹配中, 从初始姿态到最终姿态的确定包括 3 维模型的平移 (以 $\Delta a, \Delta b$ 表示) 和旋转 (以 $\Delta\varphi$ 表示)。由此可见, 匹配的关键在于寻找一组最优的参数 $(\Delta a, \Delta b, \Delta\varphi)$, 以便使得模型投影与图像的边缘信息匹配程度为最高。确定物体与模型最佳匹配的参数主要包括以下 3 个步骤: (1) 建立物体模型; (2) 建立形状匹配的优化函数; (3) 对优化函数进行求解。

4.1 车辆模型的建立

本文采用 3 维线框模型^[1]描述车辆的姿态。为了将 3 维 12 面模型投影到图像平面, 以获取模型边界线, 需要进行, 消除隐藏线以及坐标变换两方面的工作。



车辆定位系统涉及到 4 类坐标系, 即世界坐标系 (W), 模型坐标系 (M), 摄像机坐标系 (C) 以及图像坐标系 $(x - y)$ (如图 2 所示)。根据计算机视觉的有关理论^[9], 在线性摄像机模型下, 若已知模型的初始姿态以及摄像机内参数矩阵 R_1 和外参数矩阵 R_2 , 则各坐标系之间的变换关系如式 (5) 所示。

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = R_0 \cdot \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix}, \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = R_1 \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, X_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = R_2 \cdot \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} \quad (5)$$

其中

$$R_0 = \begin{bmatrix} \cos(\varphi + \Delta\varphi) & -\sin(\varphi + \Delta\varphi) & 0 & a + \Delta a \\ \sin(\varphi + \Delta\varphi) & \cos(\varphi + \Delta\varphi) & 0 & b + \Delta b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

其中, φ 为模型在世界坐标系下的初始转向角, a, b, c 分别为模型在世界坐标系下的初始坐标值。 R_1 和 R_2 由摄像机标定决定, 其定义参见文献 [9]。

可采用 CAD (computer aided design) 技术, 在 3 维摄像机坐标系下消除隐藏线, 具体方法参见文献 [10] 和 [11]。

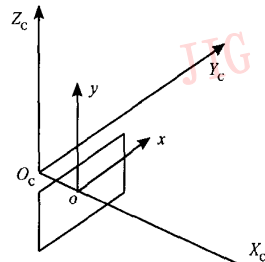


图 2 坐标系

Fig.2 Coordinate system

4.2 基于 Hausdorff 距离的匹配优化函数

设 $T = \{t_j, 1 \leq j \leq J\}$ 为从 ROI 中提取的边界轮廓点集, $M = \{m_i, 1 \leq i \leq L\}$ 为线框模型线段集合, $R = R_2 R_1 R_0$ 为从模型坐标到图像坐标的变换矩阵, 则基于 Hausdorff 距离的形状匹配可以表示如下:

$$\min(H(R(M), T)) = \min(H(\hat{M}, T)) \quad (7)$$

其中, \hat{M} 为 M 在图像坐标系下的投影, $\hat{M} = \{\hat{m}_i, 1 \leq i \leq L\}$ 。

$$\begin{aligned} H(\hat{M}, T) &= \max(h(\hat{M}, T), h(T, \hat{M})) \\ h(T, \hat{M}) &= \frac{1}{\sum_{t_j \in T} W_{t_j}} \sum_{t_j \in T} W_{t_j} \cdot \min_{\hat{m}_i \in \hat{M}} \text{mind}(t_j, \hat{m}_i) \\ h(\hat{M}, T) &= \frac{1}{\sum_{\hat{m}_i \in \hat{M}} W_{\hat{m}_i}} \sum_{\hat{m}_i \in \hat{M}} W_{\hat{m}_i} \cdot \min_{t_j \in T} \text{mind}(t_j, \hat{m}_i) \\ &= \frac{1}{L} \sum_{\hat{m}_i \in \hat{M}} \min_{t_j \in T} \text{mind}(t_j, \hat{m}_i) \end{aligned} \quad (8)$$

注:模型投影各可见线段所对应的权值均设为 1。
 其中, $d(t_j, \hat{m}_i)$ 定义为点 t_j 到线段 \hat{m}_i 的距离(如图 3 所示), 给定点 t_j 和线段 \hat{m}_i (由端点 b_1 和 b_2 确定), 则

$$d(t_j, \hat{m}_i) = \begin{cases} |t_j \hat{t}_j| & \text{若 } \overline{t_j b_1} \cdot \overline{t_j b_2} > 0 \\ \min_{b \in (b_1, b_2)} |t_j b| & \text{否则} \end{cases} \quad (9)$$

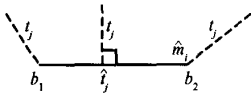


图 3 点到线段的距离

Fig. 3 Distance from point to line segment

4.3 匹配参数的求解

根据式(7)~式(9)可知, 目标函数(式(7))为连续不可微、多变量函数, 其在搜索空间中存在多个局部极小点。由于局部搜索法, 如爬山法和最陡下降法只能找到局部最优解, 因此应采取模拟退火算法^[12]或遗传算法 (genetic algorithm, GA) 来获得目标函数的全局最优解。本文实验采用模拟退火算法。

模拟退火算法的特点是简单、通用、鲁棒性强, 适用于并行处理, 其计算过程是一系列的“产生新解→计算目标函数差→判断是否满足接受准则→接受(或舍弃)新解”的迭代过程。它包含以下两个循环: 内循环表示在同一个温度 T_k 时, 在一些状态随机搜索, 它采用了基于概率的双向随机搜索技术, 当邻域的一次操作使当前解的质量提高时, 则接受这个改进解作为新的当前解, 反之, 以与 Metropolis 准则对应的转移概率来接受相对质量比较差的解作为

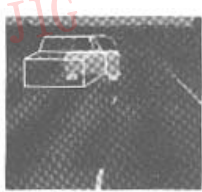
新的当前解; 外循环表示温度下降时, 系统状态的变化, 它主要包括温度下降变化 $T_{k+1} = f(T_k)$ 和算法终止条件的确定。

在通常的 SA 过程中, 算法终止于一个预先规定的停止准则, 但由于 SA 算法搜索过程的随机性, 且某些当前解要达到最优解时必须经过暂时恶化的“山脊”, 因此通常停止准则无法保证算法所得的最终解必定是最优的。本文采用带记忆的模拟退火算法, 即给算法增加一个记忆器, 使之能够记住搜索过程中遇到过的最好结果, 这样当退化过程结束时, 就可将所得最终结果与记忆器中的解进行比较, 并取较优者作为最后结果, 以便提高算法所得解的质量。记忆装置可以设置为变量 i_{best} 和 f_{best} , 其中 i_{best} 用于记忆当前遇到的最优解, f_{best} 为其目标函数值。至于记忆的实现, 开始令 i_{best} 和 f_{best} 分别等于初始解 i_0 及其目标函数值 f_0 ; 以后每接受一个新解时, 就将新当前解的目标函数值 f 与 f_{best} 进行比较, 若 f 优于 f_{best} , 则将 i 和 f 分别存入 i_{best} 和 f_{best} ; 最后算法结束时, 再从最后的当前解 i 和 i_{best} 中选取较优者为最终解。

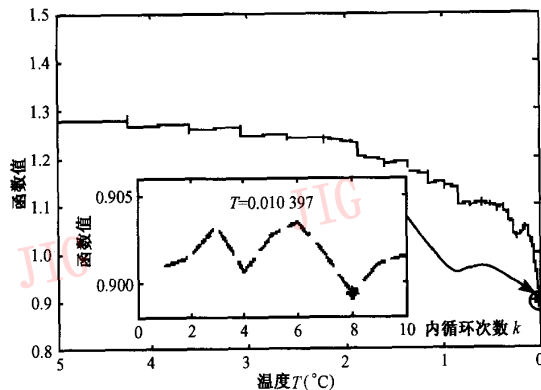
5 实验结果

为了验证上述算法的有效性, 利用 3 维仿真实验平台(以 3 维视景仿真软件 Vega 作为场景驱动, Visual C++ 6.0 为开发环境)对不同姿态的车辆图像进行了上述方法的实验, 结果分别如图 4~图 6 所示。

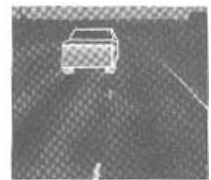
其中图 4(a)、图 5(a)、图 6(a) 表示线框模型的初始姿态, 此时的匹配函数值对应图 4(b)、图 5(b)、图 6(b) 中的初始函数值; 图 4(c)、图 5(c)、



(a) 初始姿态



(b) 姿态搜索



(c) 匹配姿态

图 4 3 维模型匹配的实验结果(姿态 1)

Fig. 4 Experimental results of 3D model matching (pose 1)

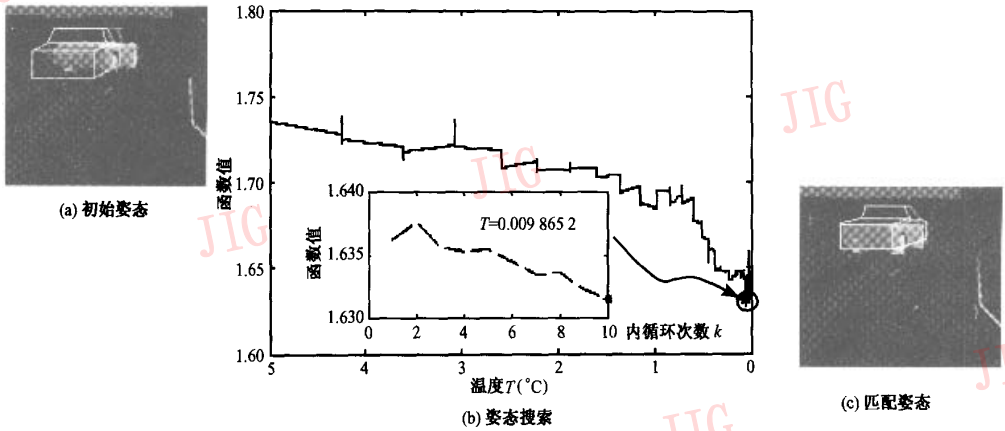


图 5 3 维模型匹配的实验结果(姿态 2)

Fig. 5 Experimental results of 3D model matching (pose2)

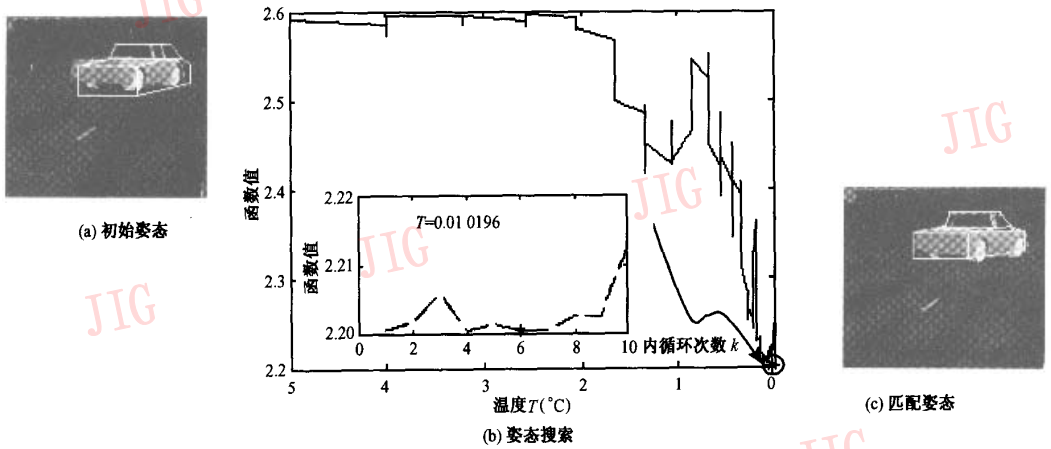


图 6 3 维模型匹配的实验结果(姿态 3)

Fig. 6 Experimental results of 3D model matching (pose3)

图 6(c) 表示线框模型与目标的最终匹配结果, 此时的匹配函数值即图 4(b)、图 5(b)、图 6(b) 中与“*”对应的函数值; 图 4(b)、图 5(b)、图 6(b) 为寻求目标函数为最小值的匹配参数的搜索过程, 图中“*”表示最优解, 其中圆圈所指向的方框内曲线图为搜索到的与最优解对应的温度下进行的内循环进程。

6 结 论

综上所述, 实现模型与图像中目标的匹配是基于 3 维模型的视觉检测与目标跟踪的首要环节。本文所提出的基于二次改进后的带权值 Hausdorff 距离的 3 维模型匹配改进方法, 与以往方法相比, 其具有以下几个特点:

(1) 以 Hausdorff 距离为基础建立模型匹配的优

化模型, 避免了传统方法中需建立模型投影与目标图像的点-点对应的过程。由于传统的距离定义对远离边界的噪声点非常敏感, 因此文中采用带有权值的改进 Hausdorff 距离作为模型投影与目标图像的匹配依据, 并根据车辆外形的特殊性, 提出了一种简便可行的目标边界提取方法, 同时以此得到的二值化边缘信息来进行图像匹配, 从而降低了后续模型匹配与搜索算法的计算代价, 并提高了匹配精度;

(2) 将模型的匹配参数定义于 3 维空间, 并通过坐标变换和距离定义来获得 3 维模型(经位移与旋转)投影与目标图像匹配函数, 由于其避免了文献[4]需要进行的在图像中搜索与 3 维模型投影中任意不在平行于 GP 平面的空间中的 3 点所对应像素点的过程, 因而降低了模型初始姿态对匹配结果的影响;

(3) 采用带记忆功能的模拟退火算法来实现匹配目标函数的全局寻优,不仅可提高搜索精度,而且算法的并行性从根本上提高了搜索效率,也使其在实际中的应用成为可能。

参考文献 (References)

- 1 Koller D, Daniilidis K, Nagel H-H. Model-based object tracking in monocular image sequences of road traffic scenes[J]. *International Journal of Computer Vision*, 1993, 10(3): 257~281.
- 2 Tan T N, Sullivan G D, Baker K D. Efficient image gradient-based vehicle localization [J]. *IEEE Transactions on Image Processing*, 2000, 9(8): 1343~1356.
- 3 Huttenlocher D P, Klanderman G A, Rucklidge W J. Comparing images using the Hausdorff distance [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1993, 15(9): 850~863.
- 4 Yang Hao, Lou J C, Sun H Z, *et al.* Efficient and robust vehicle localization[A]. In: *Proceedings of International Conference on Image Processing*[C], Thessaloniki, Greece, 2001: 355~358.
- 5 Dubuisson M P, Jain A K. A modified Hausdorff distance for object matching[A]. In: *Proceedings of 12th International Conference on Pattern recognition*[C], Jerusalem, Israel, 1994: 566~568.
- 6 Gao Y. Efficiently comparing face images using a modified Hausdorff distance[J]. *IEE Proceedings-Vision, Image and Signal Processing*, 2003, 150(6): 346~350.
- 7 Heath M D, Sarkar S, Sanocki T, *et al.* Comparison of edge detectors: a methodology and initial study [J]. *Computer Vision Image Understanding*, 1998, 69(1): 38~54.
- 8 Tan T N, Sullivan G D, Baker K D. Model-based localization and recognition of road vehicle [J]. *International Journal of Computer Vision*, 1998, 27(1): 5~25.
- 9 MA Song-de, ZHANG Zheng-You. *Computer vision—computational theory and algorithm foundation*[M]. Beijing: Science Press, 1998: 52~60. [马颂德, 张正友. 计算机视觉——计算理论与算法基础[M]. 北京: 科学出版社, 1998: 52~60.]
- 10 CHANG Ya, HU Wei-ming, TAN Tie-niu. Visualization of 3D surveillance system[J]. *Journal of English Graphics*, 2001, 22(5): 28~33. [昌娅, 胡卫明, 谭铁牛. 交通视觉监控系统中的三维车辆线框模型可视化算法[J]. *工程图学学报*, 2001, 22(5): 28~33.]
- 11 XIAO Jia. *Computer Graphics* [M]. Beijing: China Railway Publishing House, 2000: 52~64. [肖佳. 计算机绘图[M]. 北京: 中国铁道出版社, 2000: 52~64.]
- 12 KANG Li-shan, XIE Yun, YOU Shi-Yong, *et al.* Non-numerical parallel algorithm—simulated annealing algorithm [M]. Beijing: Science Press, 1994: 25~150. [康立山, 谢云, 尤矢勇等. 非数值并行算法——模拟退火算法[M]. 北京: 科学出版社, 1994: 25~150.]