

文章编号:1001-9014(2011)05-0446-05

基于多尺度特征提取的 Kalman 滤波跟踪

孔军^{1,2,3}, 汤心溢², 蒋敏¹, 刘士建², 李丹^{2,3}

(1. 江南大学 物联网工程学院, 江苏 无锡 214122; 2. 中国科学院上海技术物理研究所, 上海 200083;
3. 中国科学院研究生院, 北京 100049)

摘要: 针对波动性较大目标跟踪, 传统 Kalman 滤波算法鲁棒性和实时性不足, 提出一种基于多尺度特征提取的 Kalman 跟踪算法. 前帧目标区域特征点匹配出后续帧目标区域特征点, 并以后者特征点为中心, 建立搜索区域, 避免了遍历整幅后续帧图像, 快速地为 Kalman 滤波方程状态后验值提供了稳定的观测信号和观测残差. 实验证明, 这种作为约束条件引入传统的 Kalman 滤波方程的多尺度特征提取技术, 克服了传统 Kalman 滤波时间较长, 易发散的缺陷, 从而使其有着良好的收敛性.

关键词: 目标跟踪; 多尺度特征提取; Kalman 滤波; 收敛

中图分类号: TP391.41 **文献标识码:** A

Target tracking based on multi-scale feature extraction Kalman filter

KONG Jun^{1,2,3}, TANG Xin-Yi², JIANG Min¹, LIU Shi-Jian², LI Dan^{2,3}

(1. School of Internet of Things and Sensor Network Engineering, Jiangnan University, Wuxi 214122, China;
2. Shanghai Institute of Technical Physics, Chinese Academy of Science, Shanghai 200083, China;
3. Graduate School of Chinese Academy of Science, Beijing 100049, China)

Abstract: To overcome the deficiency of the robustness and real-time performance of traditional Kalman filter used for changeable target tracking, a new Kalman algorithm based on multi-scale feature extraction was proposed. After the feature points of a frame matched ones of the follow-up frame in the target area of image, the latter feature points centroid was took as the center from which the searching area was located so as to avoid traversing the whole image. So the stable signals and residuals of observations were provided to the Kalman filter equations to calculate accurately the posteriori state value. Experiment show that the multi-scale feature extraction technology introduced into the traditional Kalman filter equation as constrained conditions reduced filtering time and restrained divergence. Thus improved filter has good convergence.

Key words: target tracking; multi-scale feature extraction; Kalman filter; convergence

PACS: 07.05.Pj

引言

在计算机视觉研究领域, 对视频序列中目标进行跟踪已成为重要的研究方向. 针对不同的跟踪情况, 人们提出了多种目标跟踪算法^[1,2], 其中卡尔曼滤波算法是一个最优化自回归数据处理算法, 能较好且有效地解决很大部分的问题^[3]. 在利用 Kalman 滤波算法进行目标跟踪时, 观测信号的良好与否, 直接影响到最优估计结果. 背景差分和帧差分是获取目标观测信号的常用方法, 但这不仅需要每帧图像

具有良好的预处理效果, 还要对整幅图像进行全遍历, 且观测信号的精度也不高, 最终影响了 Kalman 滤波的稳健性和实时性^[4].

针对 David G. Lowe 在文献[5,6]中所提及的 SIFT 算法及其不足, 本课题在前期研究中提出了一种优化的多尺度特征提取 (Multi-Scale Feature Extraction, 简称 MSFE) 算法^[7]. 该算法通过基于尺度因子变化高斯卷积核自适应调整, 以及连续帧双向配准的办法, 来快速有效地完成连续帧图像的配准. 而本文所提及的基于 MSFE 的 Kalman 跟踪, 是在进

收稿日期: 2010-12-21, 修回日期: 2011-06-18

Received date: 2010-12-21, revised date: 2011-06-18

基金项目: 中央高校基本科研业务费专项资金 (JUSR211A36), 国家自然科学基金多年期重大国际合作研究项目 (60910005)

作者简介: 孔军 (1974-), 男, 安徽合肥人, 讲师, 中国科学院上海技术物理研究所电子科学与技术工学博士生, 主要从事机器视觉和目标跟踪方面的研究. E-mail: j-kong@163.com.

行 Kalman 滤波同时引入 MSFE 技术,前帧目标区域特征点匹配出后续帧目标区域特征点,以后者特征点为中心确定出目标在后续帧中的搜索范围,并以此作为约束条件,代入 Kalman 滤波方程,计算出当前观测值.由于该算法不仅避免对整幅图像的全遍历和预处理,而且目标区域所匹配出的特征点在背景中具有明显标定作用,从而快速地为 Kalman 滤波方程提供稳定精良的观测信号,最终提升了传统的 Kalman 滤波的稳健性和实时性.

1 MSFE 特征点提取与匹配

针对连续帧图像中波动性较大的目标会发生旋转、缩放、亮度变化特点,先构建图像的多尺度空间,再获取和优化特征点,并计算出具有位置、尺度、旋转等不变量的局部特征描述矢量.

1.1 多尺度空间建立

第 1 步:当前帧图像作为第一层,以卷积核尺寸为 10×10 模板对第一层图像进行高斯平滑处理得到第 2 层,依然用 10×10 卷积核处理第 2 层图像得到第 3 层.然后改用 12×12 模板处理第 3 层得到第 4 层,并在第四层基础上得到第 5 层、第 6 层.该帧图像生成六层平滑图像即为第一组高斯金字塔图像.

第 2 步:第 1 组高斯金字塔中首幅图像进行下采样得到第 2 组高斯金字塔图像的首幅图像,重复“第 1 步”,构建第 2 组高斯金字塔图像,以此类推,得到第 3、第 4 组等高斯金字塔图像,共 6 组.

第 3 步:对以上每组高斯金字塔图像相邻图像两两相减,得到该组的高斯差分金字塔图像,共 6 组.

第 4 步:对后续帧重复第 1 步、第 2 步、第 3 步.

1.2 多尺度空间特征点确定

去除高斯差分金字塔图像中最底层和最上层,在剩余的每一层中计算出特征点.

如果某点与该层的 8 个邻域点以及上下相邻尺度层中相对应的 18 个点的灰度值之差满足设定阈值时,该点就为兴趣点.

对该兴趣点进行泰勒级数展开^[8],取前 3 项:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (1)$$

式(1)中, $X = (x, y, \sigma)^T$ 为兴趣点的偏移量, $D = D(x, y, \sigma)$ 是兴趣点的值.对上式求导,可得 X 的极值 X_{\max} :

$$X_{\max} = - \frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (2)$$

联立式(1)和(2),可得:

$$D(X_{\max}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} X_{\max} \quad (3)$$

若 $|D(X_{\max})| \geq 0.03$,该兴趣点保留为特征点,否则,丢弃.

计算特征点相邻像素点差值,并由此形成该点的 Hessian 矩阵 H :

$$H = \begin{vmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{vmatrix} \quad (4)$$

由于 D 的主曲率和 H 的特征值成正比,令 H 特征值的最大和最小值分别为 α 和 β ,则:

$$\text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (5)$$

$$\text{Det}(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (6)$$

令 $\gamma = \alpha/\beta$,则:

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r + 1)^2}{r} \quad (7)$$

当 $r = 1$ 时, $\frac{\text{Tr}(H)}{\text{Det}(H)}$ 达到最小,并随着 r 值增加而增大,判断主曲率是否满足阈值 r 只需判断式(8)是否成立:

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(r + 1)^2}{r} \quad (8)$$

本文选取 r 经验值为 10.

用同样的方法计算出其每帧所对应的所有组高斯差分金字塔中的所有尺度上的特征点集合,而它们的交集将是该帧图像最终的特征点集.

1.3 特征点的矢量化及其向量表示

像素点的大小和方向按照如下定义:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (9)$$

$$\theta(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (10)$$

其中 $L(x, y)$ 表示 (x, y) 点处的灰度值.选取以特征点为中心的 16×16 邻域窗口,并将该窗口分割成 16 个 4×4 子窗口(每个子窗口对应一个种子),在每个子窗口中计算出所有点的大小和方向,对大小进行加权处理后可得该子窗口种子的大小.同时用直方图统计子窗口所有点在 8 个方向上的直方图.这样一个特征点由 16 个子窗口(16 个种子)的点维持着,而每个种子又可以用 8 个表示方向的数据来表示,那么最终一个特征点将由一个 128(即 16×8) 维向量来表示.

1.4 特征点的双向匹配

连续帧图像的特征向量生成之后,可用多维空间向量欧氏距离作为匹配准则,见公式(11).

$$d = \sqrt{\sum_{i=1}^{128} (x_{i1} - x_{i2})^2} \quad (i = 1, 2, \dots, 128) \quad (11)$$

其中 x_{i1} 表示其中一幅图像上的某特征点的特征向量的第 i 维元素值, x_{i2} 表示另一幅图像上的某个特征点的特征向量的第 i 维元素值,具体算法如下:

第 1 步:建立极线约束条件,减少图像配准时搜索范围;

第 2 步:求出前帧图 A 中某一个特征点 F_1 与后续帧图 B 中在对应极线上所有特征点的欧氏距离,搜索出最近点 F_2 ,如果该欧氏距离小于匹配阈值,保留 F_2 ,跳入第 3 步. 否则,删除 F_1 ,重复第 2 步;

第 3 步:求出后续帧图 B 中 F_2 与前帧图 A 中对应极线上所有特征点的欧氏距离,搜索出最近点 F_3 ;

第 4 步:判断 F_1 和 F_3 是否相等,若是,匹配成功,否则,跳入第 2 步;

2 基于 MSFE 的卡尔曼滤波

卡尔曼滤波跟踪波动性较大目标时以其位置、速度、加速度为状态矢量,通过目标的动力学方程来描述目标状态的变化,利用递推,给出目标状态的最优估计^[9,10].

卡尔曼滤波器的时间更新方程如式(12)、(13):

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} + \omega_{k-1} \quad (12)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (13)$$

定义 \hat{x}_k^- (“-”代表先验,“^”代表估计)为在第 $k-1$ 时刻对第 k 时刻的先验状态估计; \hat{x}_{k-1} 为在第 $k-1$ 时刻对 $k-1$ 时刻的后验状态估计; u_{k-1} 为第 $k-1$ 时刻的控制输入; ω_{k-1} 为第 $k-1$ 时刻的激励噪声,其协方差矩阵为 Q ; P_k^- 为在 $k-1$ 对第 k 时刻的先验估计误差协方差; P_{k-1} 为在 $k-1$ 时刻的后验估计误差协方差; A 为状态转移矩阵; B 为系统控制矩阵.

卡尔曼滤波器状态更新方程如式(14)、(15):

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (14)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (15)$$

$$P_k = (I - K_k H) P_k^- \quad (16)$$

定义 z_k 为第 k 时刻的观测值; $(z_k - H\hat{x}_k^-)$ 为观测过程残差, K_k 为 k 时刻的卡尔曼残差增益,其作用是使 k 时刻的后验估计误差协方差 P_k 达到最小; H 为

观测矩阵; R 观测噪声协方差矩阵.

卡尔曼滤波算法通过既往时刻的估计值和当前时刻的观测值,利用无偏最小均方差准则,得到当前时刻的最优估计值. 其中观测数据的准确性对估计的结果往往影响很大. 尤其当目标运动状态发生较大波动,跟踪效果并不理想. 为了提高对目标的跟踪精度和跟踪收敛速度,提出基于 MSFE 的卡尔曼滤波算法,其实质是在跟踪过程中,MSFE 算法完成相邻帧图像配准,并以每帧图像中目标区域内已经匹配成功的特征点为中心,建立相应搜索窗口,使得目标在背景不断波动时,搜索窗口做自动调整,从而快速且准确地得到目标(观测信号)的面积、质心和外接区域等参数. 同时,利用卡尔曼滤波不断地预测、修正的递推过程,最终得到运动目标的状态的最优估计.

3 实验结果与分析

为了验证本文提出的基于 MSFE 的卡尔曼跟踪算法,利用 HIROBO Shuttle SCEADU Evolution 50 型无人机自带的单目摄像机采集视频图像序列,以完成对一个“做自由落体运动的小球”进行跟踪,且小球落地后反弹,直到静止,图像分辨率为 320×240 . 系统采样时间为 $dt = 1$ s,重力加速度 $g = 9.8$ m/s²,持续时间为 60 s,运行机器: Intel Core 2 Duo E4700 2.6 GHz PC,内存 2 GB,操作系统 WinXP,开发平台 VC++6.0.

卡尔曼滤波参数如下,可假设过程激励噪声和观测噪声协方差为恒定值.

$$A = \begin{bmatrix} [1, 0, 0, 0]^T, [0, 1, 0, 0]^T, [dt, 0, 1, 0]^T, \\ [0, dt, 0, 1]^T \end{bmatrix}$$

$$B * u = [0, 0, 0, g * dt]^T$$

$$H = \begin{bmatrix} [1, 0]^T, [0, 1]^T, [0, 0]^T, [0, 0]^T \end{bmatrix}$$

$$Q = \begin{bmatrix} [0.01, 0, 0, 0]^T, [0, 0.01, 0, 0]^T, \\ [0, 0, 0.01, 0]^T, [0, 0, 0, 0.01]^T \end{bmatrix}$$

$$R = \begin{bmatrix} [0.2845, 0.0045]^T, [0.0045, 0.0045]^T \end{bmatrix}$$

卡尔曼滤波启动初始化为:

\hat{x}_0 为 100×4 的零矩阵;

$$P_0 = \begin{bmatrix} [100, 0, 0, 0]^T, [0, 100, 0, 0]^T, [0, 0, 100, 0]^T, [0, 0, 0, 100]^T \end{bmatrix};$$

截取视频流中的第 12 ~ 第 17 帧图像,按照上文所提及的多尺度特征提取法计算出每帧图像的特征点矢量集合,见图 1. 其中箭头起点代表特征点位置,箭头方向代表该特征点所处邻域的主梯度方向,箭头长度代表矢量大小.

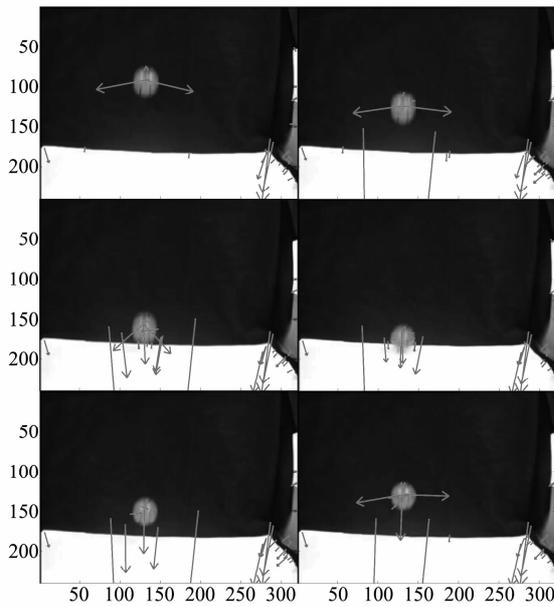


图 1 连续帧图像特征点的梯度矢量
Fig.1 Gradient of feature points in sequential frame image based on time-axis

然后以匹配阈值为 0.53,对视频流相邻帧以 MSFE 算法完成两两配准,也就是帧 1 配准帧 2,帧 2 配准帧 3 等等.由于篇幅限制,本文只给出第 12 帧与第 13 帧,第 13 帧与第 14 帧,第 14 帧与第 15 帧,第 15 帧与第 16 帧图像配准结果,如图 2,其余同理.

从图 2 看出,利用 MSFE 算法对相邻帧匹配效果很好,能准确地标定出相邻帧之间的目标区域(小球)上特征点的匹配关系,这为后续的卡尔曼滤波时避免遍历整幅图像而直接搜索目标区域奠定基础.另外,虽然目标区域外也存在着特征点匹配关系,但由于其数量有限,并不影响卡尔曼滤波结果.

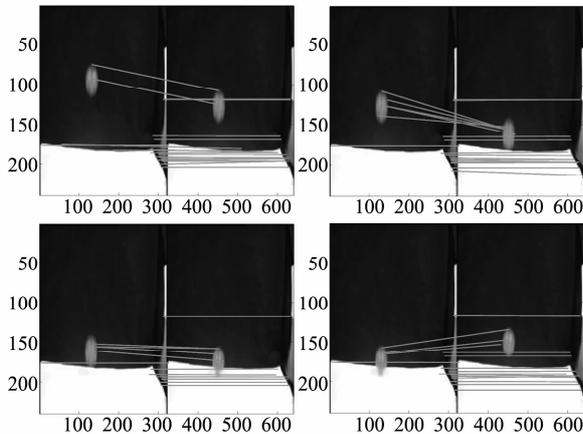


图 2 基于 MSFE 算法的特征点匹配结果(匹配阈值 = 0.53)
Fig.2 Result of matching based on MSFE algorithm (threshold = 0.53)

预先设定状态和误差协方差的初始化值 \hat{x}_0 和 P_0 ,整个卡尔曼滤波过程即可启动并持续递推下去.每帧图像中目标区域所配准出的特征点,将作为卡尔曼滤波过程中每步的约束条件,优化卡尔曼滤波算法,以抵消目标在背景中的大范围变化时所带来的观测信号的波动误差.同时,以目标区域的特征点为中心,进行邻域搜索,确定目标位置,快速地为卡尔曼滤波方程提供观测信号.图 3 中,小球外的浅色圆环,即是卡尔曼跟踪的估计结果,从图中可得,估计值与观测值非常接近,滤波效果良好.

图 4 是针对试验中“自由落体且落地反弹小球”基于 MSFE 的 Kalman 滤波所得状态最优估计,不难看出,从第 6 帧开始,算法在水平方向能准确地估计出自由落体小球位置基本没有变化,同时看出,从第 50 帧开始,算法在垂直方向也能准确地估计出小球趋于稳定,滤波得到很好的收敛^[11].

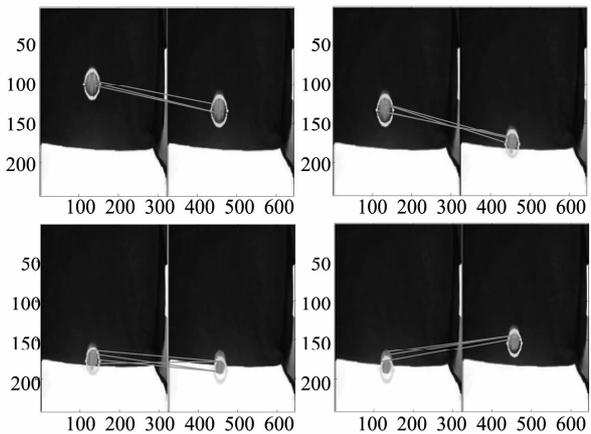


图 3 基于 MSFE 的卡尔曼滤波结果
Fig.3 Result of Kalman filter based on multi-scale feature extraction algorithm

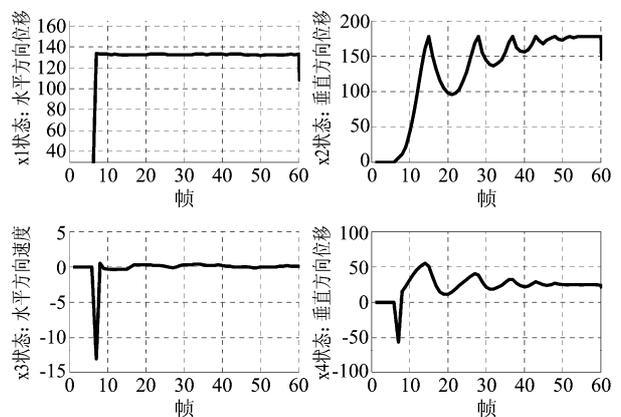


图 4 基于 MSFE 的卡尔曼滤波状态分析
Fig.4 States of Kalman filter based on multi-scale feature extraction algorithm

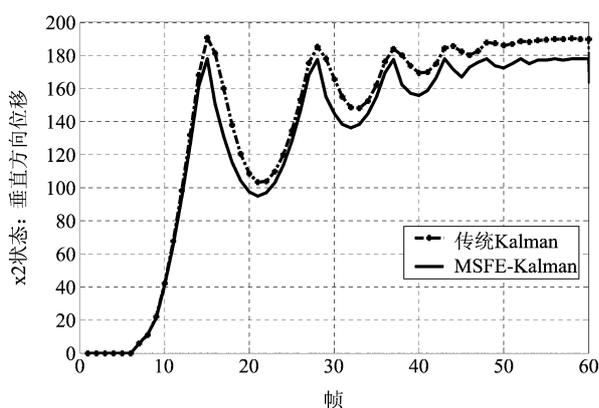


图5 Kalman 和 MSFE-Kalman 垂直方向位置状态对比
Fig. 5 The vertical position's state contrast between Kalman and MSFE-Kalman

为了进一步说明本文所描述滤波的优点,现将 MSFE-Kalman 算法和传统 Kalman 算法在对小球的垂直方向位置状态估计进行对比,如图 5. 在目标状态估计时,由于引入 MSFE 图像配准作为滤波的约束条件,所以得到观测信号相比传统算法精确度要高,而且获取时间也较短. 从图中对比可看出,实线的峰值较虚线的峰值超前,说明 MSFE-Kalman 滤波算法时效性强;另外实线的峰值几乎都稳定在同一个峰值上(本次实验稳定值为 177.833 左右),说明 MSFE-Kalman 滤波算法的收敛性很高.

4 结语

本文针对波动性较大的运动目标跟踪问题,提出了一种基于多尺度特征点提取(MSFE)的 Kalman 滤波的跟踪算法. 该算法将 MSFE 特征点配准作为约束条件引入 Kalman 滤波方程,以前帧图像中目标所匹配出的后续帧图像目标区域上的特征点为中心,建立目标小范围搜索区域,避免对整幅图像进行遍历的传统算法,同时所匹配出的特征点具有目标位置标定作用,为当前时刻状态后验估计提供稳定

且精确的观测值和观测残差,从而提高 Kalman 滤波的鲁棒性和时效性. 实验证明,基于 MSFE 的 Kalman 滤波算法在跟踪波动性较大的运动目标,其状态各个分量都能得到快速且有效地收敛.

REFERENCES

- [1] Bakowki A, Jones G A. Video surveillance tracking using colour region adjacency graphs[J]. *IEEE Conference Publication*. 1999,2(465):794-798.
- [2] Wixson L. Detecting salient motion by accumulating directionally-consistent flow[J]. *IEEE Transactions of Pattern Analysis and Machine Intelligence*. 2000,22(8):774-780.
- [3] XIN Yun-Hong, YANG Wan-Hai. Pseudo-linear kalman filter with application to the passive tracking of multi-stations IRSTS[J]. *J. Infrared Millim. Waves*(辛云宏,杨万海. 基于伪线性卡尔曼滤波的多站 IRST 系统跟踪技术. *红外与毫米波学报*),2005,24(5):374-377.
- [4] A Shahbahrami, B Juurlink, S Vassiliadis. Accelerating color space conversion using Extended subwords and the matrix register file[C]. *Eighth IEEE International Symposium on Multimedia*,2006:37-46.
- [5] David G. Lowe. Object recognition from Local scale-invariant features[J]. *International Conference on Computer Vision, Corfu, September,1999*,3(1):1150-1157.
- [6] David G. Lowe. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*. 2004. 60(2):91-110.
- [7] Kong Jun, Tang Xin-Yi, Jiang Min. Object location technique for moving target based on multi-scale feature extraction[J]. *J. Infrared Millim. Waves*(孔军,汤心溢,蒋敏. 基于多尺度特征提取的运动目标定位研究. *红外与毫米波学报*),2011,30(1):21-26.
- [8] WANG Hong-Bing, PENG Zhen-Ming, LIU Jie, et al. Feature points detection and tracking based on SIFT combining with KLT method[J]. *Proc. SPIE*. 2009(7506):75062N1-10.
- [9] R E Kalman, A new approach to linear filtering and prediction problem[J]. *Trans on ASME. Journal of Basic Engineering*,1960,82(1):35-45.
- [10] Kalman R E, Bucy R S. New methods and results in linear filtering and prediction theory[J]. *Trans on ASME, Journal of Basic Engineering*,1961,83:95-108.
- [11] Meyn S P, Lyndon J B, Model reference adaptive control of time varying and stochastic systems[J]. *IEEE Trans on Auto Control*,1993,38(2):1739-1753.

(上接 389 页)

- [6] YU Wei-Hua, MOU Jin-Chao, LI Xiang, et al. Design of a W band Subharmonically pumped Mixer Based on a hybrid Schottky Diode Model[C]. *2009 International Conference on Microwave Technology and Computational Electromagnetics*,2009:207-210.
- [7] YANG Zi-Qiang, YANG Tao, LIU Yu. Design of Ka-band monolithic image rejection mixer[J]. *International Journal of Infrared and Millimeter Waves*,2007,28(3):237-241.
- [8] MOU Jin-Chao, YUAN Yong, YU Wei-Hua, et al. Design and Fabrication of planar GaAs Schottky Barrier Diodes for Submillimeter-wave Applications[C]. *International Confer-*

ence on Microwave and Millimeterwave Technology, 2010: 1746-1749.

- [9] RHODERICK E H. Metal-semiconductor contacts[C]. *IEE Proceedings-Communications*,1982,129(1):1-14.
- [10] ENGEN G F, HOER C A. Thru-Reflect-Line: An Improved Technique for Calibrating the Dual Six-Port Automatic Network Analyzer[J]. *IEEE Transactions on Microwave Theory and Techniques*,1979,27(12):987-993.
- [11] WANG Chuang, GU Jianzhong, SUN Xiaowei. Low cost compacted Ka-band rat-race mixer[C]. *Asia-Pacific Conference Proceedings*,2005.