

基于仿射不变约束与快速 EKF 滤波的航拍图像稳像

易 盟 楚 岩

(长安大学电子与控制工程学院 西安 710064)

摘 要 考虑到航拍机载成像平台抖动严重、视频稳像匹配环节精度不一致的特点以及航拍图像稳像技术快速、准确的要求,提出了一种结合仿射不变约束与快速扩展卡尔曼(Extend Kalman Filter, EKF)滤波的图像稳像算法。该算法首先以视频参考帧中的角点量作为特征点,通过 Harris 检测器选择出稳定角点;然后对待配准点构建 Delaunay 三角网进行初始匹配,提出利用仿射不变约束方法筛选出精确匹配点;最后利用快速 EKF 运动滤波方法实时估计和修正噪声的统计特性,从而解决摄像机扫描运动中存在的抖动问题。在对大量分辨率为 640×480 pixel 的航拍图像的仿真实验中,可通过仿射不变约束实现精确的模型估计,采用的快速运动补偿方法在补偿过程中耗时为 5.054ms,比传统的运动补偿方法节约了 69.5% 的时间。实验结果表明,该算法能够实时稳定航拍视频帧间的抖动现象,并能有效跟随场景的真实扫描。

关键词 航拍图像,电子稳像,特征点,仿射不变约束,EKF 运动滤波

中图分类号 TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.8.064

Aerial Video Image Stabilization Based on Affine Invariant Constraint and Fast EKF Adaptive Filter

YI Meng CHU Yan

(School of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China)

Abstract In view of serious jitter of aerial airborne imaging, different accuracies of obtained matching points in aerial video stabilization, and requirements of fast and precise aerial image stabilization technology, an image stabilization algorithm was proposed combining affine invariant constraint and Extend Kalman Filter (EKF) adaptive filter. Firstly, corners are detected from reference frame as feature points, the locally stable points are selected by the Harris detector. Then Delaunay triangulation is used to find initial matching, and the most accurate matching points which best satisfy the affine invariant constraint are filtered. Finally, EKF motion filter method is used to estimate and correct the statistical property of noise, so jitter of the camera can be removed but scanning motion can be retained simultaneously. As to a large number of simulation experiment of aerial images with a resolution of $640 \text{ pixel} \times 480 \text{ pixel}$, accurate model can be estimated by affine invariant constraints, the rapid motion compensation method takes 5.054 ms in the process of compensation, and saves for 69.5% than traditional motion compensation method. The experimental results illustrate that the proposed algorithm can stabilize the inter-frame jitter and track the real scene effectively.

Keywords Aerial video, Electronic image stabilization, Feature point, Affine invariant constraint, EKF motion filter

1 引言

电子稳像(Electronic Image Stabilization)技术是应用图像处理 and 电子技术的方法来直接确定视频帧间偏移量并进行补偿的技术。目前电子稳像技术已广泛应用于目标检测与跟踪、步行机器人、视频压缩和图像拼接等领域^[1,2]。航拍图像基于航空成像平台获取对地观测图像,机载成像平台的随机运动、抖动等所造成的图像序列的不平稳性,使得如何针对航拍图像的特点设计合适的电子稳像算法成为提高航空监控技术的关键。

电子稳像系统主要由运动估计和运动补偿两部分构成。运动估计是根据图像序列的各种信息找出相邻帧之间的运动

偏移量,其主要算法有块匹配法、灰度投影法、位平面法及特征匹配法等,其中基于特征匹配的方法^[3-5]获得了最广泛的应用。直接匹配法^[3]通过直接比较特征描述符获取一致匹配点,但该方法会产生大量的错误匹配对,从而导致运动估计不准确。Yeong 等^[4]利用一对一的匹配抑制方法将匹配问题转化为优化问题。Chen 等^[5]利用空间排序抑制方法保留临近点并剔除可疑点。但是上述匹配方法的效果都接近基于特征描述符方法的效果,也就是说,在大的尺度变化、旋转、光照和视角变化时,利用上述方法得到的很多正确匹配对会被丢弃。运动补偿是通过前一帧局部图像的运动信息来预测、补偿当前帧的局部图像的过程。常用的补偿方法有衰减因子法^[6]、均匀滤波法^[7]、卡尔曼滤波法^[8,9]等。其中,衰减因子法中的

到稿日期:2015-12-07 返修日期:2016-03-16 本文受国家自然科学基金(41372330),中央高校基本科研业务费专项资金(310832151097)资助。

易 盟(1982-),男,博士,讲师,主要研究方向为模式识别、图像处理等, E-mail: yimeng0120@163.com; 楚 岩(1962-),女,博士,教授,主要研究方向为虚拟现实、计算机视觉等。

衰减系数需要人工预先设置,因此无法自适应于航拍视频序列;均匀滤波为典型的线性滤波算法,用模板像素的平均值代替原图像的像素值,但该方法不能很好地保留图像细节且不能去除噪声点。卡尔曼滤波通过利用前一时刻的估计值来更新现在时刻的状态变量,并求出现在时刻的估计值,但该方法存在滤波次优甚至发散的可能。

针对上述航拍视频图像在精度和稳定性方面面临的问题,在对运动估计和运动补偿各个环节进行详细分析的基础上,提出一种基于仿射不变约束与快速扩展卡尔曼滤波的航拍图像稳像算法。该算法首先采用 Harris 算子得到参考帧中的特征点,利用 Delaunay 三角网得到其在当前帧中的匹配点,然后利用仿射不变约束提取具有高精度的匹配点,再将验证后的正确匹配特征点对代入运动方程,通过求解最小二乘解来获得运动估计矢量。在运动补偿方面,提出利用 EKF 运动滤波方法对 Kalman 滤波中的过程噪声和观测噪声进行在线估计,得到最终的补偿矢量,同时利用图像旋转的快速补偿算法提高运动补偿的效率,实现实时航拍图像稳像。

2 图像运动参数模型

假设运动的摄像机对着一个平坦的面进行拍摄,并且摄像机的光轴与该平面垂直,此时摄像机的平移运动就会导致 2 维图像以均一比例变化和位移,摄像机垂直轴的转动导致图像的整体旋转。由此建立 2 维仿射模型,如式(1)所示。

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix} \quad (1)$$

其中, (X_{t-1}, Y_{t-1}) 和 (X_t, Y_t) 分别为 $t-1$ 和 t 时刻二维平面中的坐标, θ 是旋转角度, $(\Delta X, \Delta Y)$ 为水平和垂直方向上的平移量。当旋转角度较小(小于 5°)时, $\cos\theta \approx 1$, $\sin\theta \approx \theta$, 并以此线性化原仿射模型式(1)得到相似变换模型,如式(2)所示。

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} 1 & -\theta \\ \theta & 1 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix} \quad (2)$$

3 全局运动估计

3.1 特征点提取

航拍视频得到的城市场景图像中存在大量的角点特征,因此,如何快速准确地提取出参考帧中的角点至关重要。Harris 检测器在特征点检测稳定性方面具有优良的性能^[10],因此本文将利用 Harris 检测器检测得到的点作为特征点。Harris 特征点检测步骤如下:

$$M = \omega \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (3)$$

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 \quad (4)$$

其中, ω 为高斯函数; I_x, I_y 分别为图像在水平和垂直方向上的梯度; λ_1, λ_2 为自相关矩阵 M 的特征值; R 为兴趣值,通过计算 R 得到图像特征点。

3.2 特征点匹配

本文利用待配准图像的 Delanay 三角网对两幅具有几何相似性的航拍图像进行特征匹配^[11,12]。虽然利用 D-三角网法可以得到特征点的初始匹配结果,但是由于航拍图像分辨率较高,计算比较耗时,因此本文在特征点匹配前对特征点进行筛选,以减小一致特征点匹配的计算量。首先,选定最长的 200 条 Delanay 三角网的边进行匹配,保证得到更稳定和准确

的匹配结果;其次,由于两个航拍图像帧间的视角差不大,两帧图像一致区域的 Delanay 三角网边方向应该相对一致,因此,本文只对角度差小于 10° 的两条边进行特征匹配。最后,利用 Delanay 三角网边两侧局部区域的像素强弱的对比结果,可以提前筛选出一定的伪匹配点。假设在 Delanay 三角网边的一侧的平均像素强度响应值为 l_1 ,而在 Delanay 三角网的另一侧的平均像素强度响应值为 l_2 。令像素强度对比值 S 为 $l_2 - l_1$,当 $l_2 - l_1 \geq 0$ 时, S 取 1;当 $l_2 - l_1 < 0$ 时, S 取 -1。因此,如果匹配前两条边的强度对比值 S 不相等,则该待匹配对被排除。图 1 分别给出了对两幅待稳像图像的特征点构造的 Delaunay 三角网图,可以看出两幅图像具有很多相同的 Delaunay 三角网结构。



图 1 两幅航拍图像分别由 40 个稳定角点生成 Delaunay 三角网

3.3 基于仿射不变约束的精准匹配对提取

由齐次坐标系下的 3 个点 $p_1 = (x_1, y_1, 1)$, $p_2 = (x_2, y_2, 1)$, $p_3 = (x_3, y_3, 1)$ 组成的三角形的净面积为:

$$TAB(p_1, p_2, p_3) = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \quad (5)$$

假设两组一致匹配点对集分别为 $P' = (P_1, P_2, \dots, P_m)$, $Q' = (Q_1, Q_2, \dots, Q_m)$, 则对任意一对匹配点 (P_i, Q_i) , 根据与 P_i 的欧氏距离,在 P' 中选择出 k 近邻(kNN) ($k=5$) 点集,其中 m 为两幅图像得到的匹配对数, i 为匹配对数中的任意一项,且其取值范围为 $1 \leq i \leq m$ 。定义离 P_i 最近的点为起始点,并根据逆时针角度对该点集进行排序,将该序列定义为 $P_i = \{P_{i1}, P_{i2}, \dots, P_{ik}\}$, 将对应于点集 P_i 的匹配点序列 Q' 定义为 $Q_i = \{Q_{i1}, Q_{i2}, \dots, Q_{ik}\}$ 。

对于任意点 (P_i, Q_i) , 定义 $s_{xy} = |TAB(P_i, P_x, P_y)|$, 其中 $|\cdot|$ 表示绝对值, $x=1, 2, \dots, k, y=1, 2, \dots, k, x \neq y$, 同样对应的匹配点集有 $s'_{xy} = |TAB(Q_i, Q_x, Q_y)|$ 。如果 (P_x, Q_x) 和 (P_y, Q_y) 为正确的匹配对,则能够得到:

$$\frac{s_{12}}{s_{13}} \approx \frac{s_{23}}{s_{23}} \approx \dots \approx \frac{s_{k1}}{s_{k1}} \quad (6)$$

定义 $TAB_i = \{s_{12}/s'_{12}, s_{23}/s'_{23}, \dots, s_{k1}/s'_{k1}\}$, 这样,将 TAB_i 的方差作为点集 (P_i, Q_i) 的仿射不变空间约束项,并定义为 $AISC(P_i, Q_i) = Var(TAB_i)$, 当 $AISC(P_i, Q_i) \leq \xi$ 时,则认为 (P_i, Q_i) 为正确匹配对,其中, ξ 为给定的阈值。

将通过仿射不变约束得到的 N 对匹配点代入式(2),得到含有 3 个未知参数的 $2N$ 个线性方程。对方程进行展开、移项,整理得到 $B=Am$, 其中:

$$B = \begin{bmatrix} x_{t,1} - x_{t-1,1} \\ y_{t,1} - y_{t-1,1} \\ \vdots \\ x_{t,N} - x_{t-1,N} \\ y_{t,N} - y_{t-1,N} \end{bmatrix}, A = \begin{bmatrix} -y_{t-1,1} & 1 & 0 \\ x_{t-1,1} & 0 & 1 \\ \vdots & \vdots & \vdots \\ -y_{t-1,N} & 1 & 0 \\ x_{t-1,N} & 0 & 1 \end{bmatrix} \quad (7)$$

$$m = [\theta, \Delta X, \Delta Y]^T$$

根据解超定线性方程组的方法求解 m , m 的最小二乘解由 $m = (A^T A)^{-1} A^T B$ 得到。

4 帧间运动矢量补偿

当摄像机存在跟踪拍摄时,摄像机的运动由随机抖动和正常扫描运动构成。对当前帧进行补偿时须明确:待补偿的矢量仅是摄像机的随机抖动,而扫描运动是需要保留的矢量。只有这样,才能在保留跟踪拍摄的同时去除或减轻摄像机的抖动。卡尔曼滤波法基于这样的假设:摄像机的扫描运动变化平稳,抖动的变化随机性大。因此利用卡尔曼滤波首先得到平滑运动分量(x_{filter}),而最终的抖动分量则是原始运动矢量(x_{mov})与平滑运动分量的差值,即 $x_{jitter} = x_{filter} - x_{mov}$ 。

4.1 Kalman 滤波

Kalman 滤波是一种应用较广的动态数据处理方法,但在应用时要求模型参数和噪声的先验已知,在许多条件下它们是未知的或近似已知的,这将导致滤波精度降低甚至发散。EKF 滤波^[13]在计算过程中利用得到的数据进行递推滤波的同时,实时地对不确切的模型参数和噪声统计特性进行适当估计和优化,这样可以更好地实现精确性。EKF 滤波模型如下:

$$\begin{cases} S(k|k-1) = F \times S(k-1) \\ P(k|k-1) = F \times P(k-1) F^T + \hat{Q}(k-1) \\ S(k|k) = s(k|k-1) + K(k) \times \epsilon(k) \\ \epsilon(k) = Z(k) - H \times S(k|k-1) \\ K(k) = P(k|k-1) \times H^T (H \times P(k|k-1) \times H^T + \hat{R}(k))^{-1} \\ P(k|k) = (I - K(k) \times H) \times P(k|k-1) \end{cases} \quad (8)$$

其中, F 为状态变换模型, H 表示测量模型, R 为等价测量噪声矩阵, Q 为过程噪声, $P(k-1)$ 为预状态模型, $P(k|k-1)$ 为状态预测模型。

$\hat{R}(k), \hat{Q}(k)$ 表示为:

$$\hat{R}(k) = (1 - d_k) \hat{R}(k-1) + d_k (\epsilon(k) * \epsilon(k)^T - H * P(k) * H^T) \quad (9)$$

$$\hat{Q}(k) = (1 - d(k)) \hat{Q}(k-1) + d(k) (K(k) * \epsilon(k)^T K(k)^T + P(k|k) - F * P(k-1|k-1) * F^T) \quad (10)$$

其中, $d(k) = (1 - b) / (1 - b^k)$, b 为遗忘因子且 $0 < b < 1$ 。

4.2 EKF 滤波算法的改进

为了避免不稳定的滤波结果,对 $P(k|k-1)$ 进行优化,通过利用预测残差对动态向量进行修正,计算出更接近实际的状态向量。首先利用信息序列^[14]来处理滤波模型的发散问题:

$$\epsilon(k)^T * \epsilon(k) \leq \gamma * \text{Trace}(H * P(k|k-1) * H^T + \hat{R}(k)) \quad (11)$$

其中, $\gamma > 1$ 。若式(11)成立,则处于正常工作状态;若不成立,说明误差过大,也就是大于估计值的 γ 倍,这样得到的结果不稳定。本文在 $P(k|k-1)$ 中加入加权系数 $C(k)$ ^[15] 并进行实时优化:

$$P(k|k-1) = C(k) * H(k, k-1) * P(k-1) * H^T(k, k-1) + R(k) Q(k) R^T(k) \quad (12)$$

其中, $C(k) \geq 1$ 。

将上式结合滤波系统模型,可以得到运动补偿向量。

4.3 快速运动补偿方法

传统的运动补偿方法是将其 $m = [\theta, \Delta X, \Delta Y]^T$ 代入式(2)进行逐点运算,其运算量大,执行效率低。本文基于图像的结构,利用像素点间的相对位置关系在图像旋转时不发生变化的特点,提出了一种图像旋转的快速补偿算法,从而大大提高了运动补偿的效率。

如图 2(a) 所示,图像旋转前,任意像素点 $E(x, y)$, $E1(x_1, y_1)$, $E2(x_2, y_2)$ 及 $A(x_A, y_A)$ 在几何上是矩形的 4 个顶点。由于旋转变换是线性变换,如图 2(b) 所示,图像旋转后,任意像素点之间的相对位置关系不发生变化,其对应点的坐标关系满足:

$$\begin{cases} x' = x_1' + (x_2' - x_1') \\ y' = y_1' + (y_2' - y_1') \end{cases} \quad (13)$$

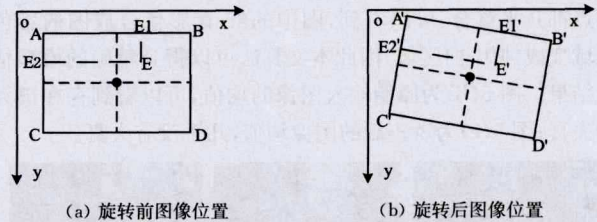


图 2 图像旋转位置关系

因此,利用首行首列像素的旋转结果就能计算出其它所有像素的旋转结果。具体步骤为:对第一行和第一列的像素通过式(2)作坐标变换;对其余行列的像素用式(13)进行简单的加减运算。这样避免了对整幅图像的每个像素点作矩阵乘法运算,从而有效地节省了运算时间。

5 实验结果及分析

为了验证本文算法在图像配准方面的有效性,对大量航拍实验图像进行了测试。实验采用分辨率为 640×480 的航拍视频序列,序列的总帧数为 80。该序列存在水平和垂直方向的随机抖动以及旋转运动,同时摄像机在水平方向上存在扫描运动。

5.1 特征检测及图像匹配

为了验证本文算法在图像检测与图像匹配方面的有效性,选用图 3 所示的典型航拍图像作为测试图像,该图片来自 DARPA 航拍视频数据,图像分辨率为 640×480 。利用红-青色合成图来说明它们之间的像素差,从图 3(c) 可以看到帧 (a) 和 (b) 之间有很明显的水平方向和垂直方向的像素偏差。图 4 为本算法对 DARPA 航拍图像进行特征点检测和匹配的结果,图 4(a) 和 (b) 中“+”为使用分区 Harris 检测器提取的兴趣点,两幅图像兴趣点数量分别为 147 和 129。图 4(c) 为两幅图像的一致匹配对分布图,正确的兴趣点匹配对为 42 个。

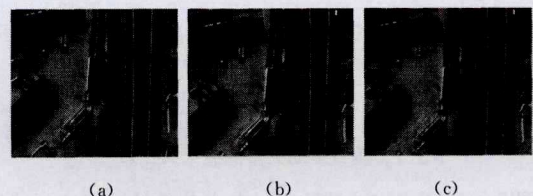


图 3 待配准航拍图像



(a)和(b)为图像检测结果;(c)为图像匹配结果

图4 图像检测和图像匹配结果

5.2 运动模型估计

实现了特征匹配后,就可以对两帧图像进行运动模型参数估计,需要由正确匹配对得到映射关系,通过变换矩阵对输入图像进行几何变换并配准。图5(a)为对输入图像帧经过模型变换并叠加在参考图像帧的彩色合成图,其结果显示,匹配点都几乎重合,可以看到,图中的红-青彩色合成图的彩色区域变成纯黑-白区域,因此本文算法可以得到精确的模型估计结果。图5(b)为原始输入图像的均值,可以看到存在很大的失真;图5(c)为矫正后的图像均值,几乎没有失真。



(a)经过几何变换后的彩色合成图 (b)原始输入图像均值 (c)矫正输入图像均值

图5 运动模型估计结果

5.3 运动矢量滤波结果比较

为了对运动矢量滤波进行测试,选用图6所示的2组航拍视频序列作为测试库,测试数据分别来自 DARPA 航拍视频、卡耐基梅隆大学 VSAM 无人机数据库和自建航拍视频数据库,每组序列各取 25 帧,分辨率为 640×480 。

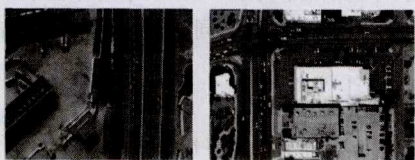


图6 2组航拍图像序列

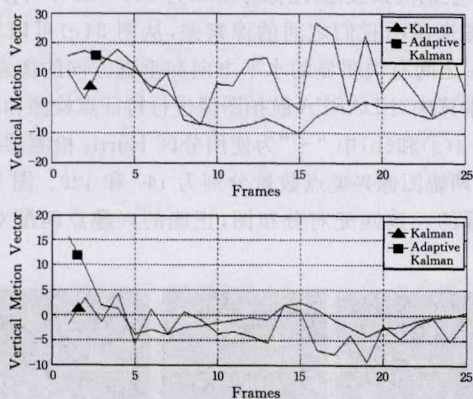


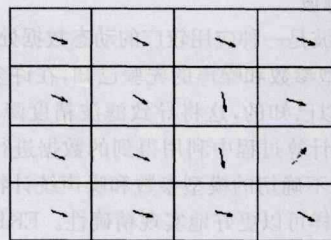
图7 垂直运动矢量滤波结果比较

图7给出了本文滤波法和经典卡尔曼滤波方法的滤波结果,可以看到由于摄像机平台的抖动造成垂直方向存在较大范围的波动。由实验结果可见,卡尔曼滤波结果偏离了航拍

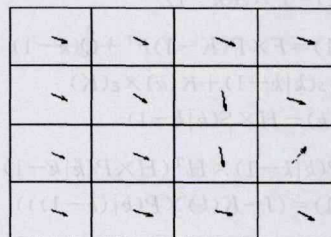
相机的运动方向,从而造成滤波结果的不稳定;而使用本文的滤波法能够很好地平滑抖动分量,并能有效地跟踪航拍相机的实际运动。

5.4 运动矢量补偿比较

图8显示的是直接对图3所示航拍序列快速运动补偿化的运动向量图结果。通过估计得到的初始运动估计向量为 $s = [0.35, 9.83, -5.26]$,通过补偿后得到的该向量为 $s = [0.12, 3.24, -1.63]$,可以看到补偿后各区域运动较为一致。初始运动向量标准差: $\sigma_x = 10.24, \sigma_y = 5.34$;补偿后的运动向量标准差: $\sigma_x' = 5.75, \sigma_y' = 3.81$ 。因此减小的标准差使运动向量更趋于一致。同时,本文的快速运动补偿方法的耗时为 5.054ms ,仅为传统的补偿方法耗时 16.611ms 的 69.5% 。



(a)补偿前的运动矢量



(b)补偿后的运动矢量

图8 分区运动矢量比较

结束语 精确、稳定的航拍视频稳像是航空电子侦察系统的核心和关键技术。本文在改进航拍视频稳像算法的基础上,提出了结合仿射不变约束与快速 EKF 滤波的图像稳像算法。本文算法通过 4 部分实验验证了图像配准的有效性:第一部分利用 DARPA 航拍视频数据验证本文算法在图像检测与图像匹配方面的有效性,结果显示本文方法能够准确地提取特征点并去除误匹配;第二部分利用仿射不变约束实现帧间图像配准,根据实验,此方法有效解决了帧间配准的失真现象;第三、四部分通过快速 EKF 滤波补偿连续帧的随机抖动现象,相对于卡尔曼滤波,本方法实现了不依赖预先设定的噪声方差阵对滤波结果的影响。

结合实验,可得到以下结论:文中方法稳定效果好,具有较强的鲁棒性,可以有效稳定航拍视频的抖动现象并跟随场景的真实扫描。尽管本文取得了较好的结果,但对于存在更为复杂的扫描运动,仍需作进一步的研究与探讨。

参考文献

- [1] Ratakonda K. Real-time digital video stabilization for multimedia applications[C]//IEEE International Symposium on Circuits and Systems. 1998:69-72
- [2] Aysun A, Ertürk S. Fast digital image stabilization using one bit transform based sub-image motion estimation[J]. IEEE Transactions on Consumer Electronics, 2005, 51(3): 917-921

- [3] Byrne J. Nested motion descriptors[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2015;502-510
- [4] Yeong K,Chulwoo L,Chang K. Video stabilization based on feature trajectory augmentation and selection and robust mesh grid warping[J]. IEEE Transactions on Image Processing, 2015, 24(12):5260-5273
- [5] Chen Y,Lin H. Full-frame video stabilization via SIFT feature matching[C]//2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2014;361-364
- [6] Sung K, Lee S, Seung J. Fast digital image stabilizer based on gray-coded bit-plane matching[J]. IEEE Transactions on Consumer Electronics, 2009, 45(3):598-603
- [7] Shih Y Y,Su S F,Rudas I. Fuzzy based hand-shake compensation for image stabilization[C]//2012 International Conference on System Science and Engineering, 2012;1511-1521
- [8] Jiang J G, Niu J J, Qi M B. Rotational video stabilization algorithm based on SURF and trajectory smoothing[J]. Chinese Journal of Scientific Instrument, 2014, 24(13):743-748(in Chinese)
蒋建国,牛杰杰,齐美彬.基于SURF和轨迹滤波的旋转视频稳像算法[J].仪器仪表学报,2014,24(13):743-748
- [9] Zhu J, Guo B. Electronic image stabilization system based on global feature tracking[J]. Journal of Systems Engineering and Electronics, 2008, 19(2):228-233
- [10] Dong K, Hong R. New image quality metric using the Harris response[J]. IEEE Signal Processing Letters, 2009, 16(7):616-619
- [11] Ming Z. A robust Delaunay triangulation matching for multi-spectral/multidate remote sensing image registration[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(4):711-715
- [12] Hongsheng L. Object matching using a locally affine invariant and linear programming techniques. [J] IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(2):411-424
- [13] Houtekamer P, Herschel L. Ensemble kalman filtering [J]. Quarterly Journal of the Royal Meteorological Society, 2006, 131(613):3269-3289
- [14] Lynn R, Alireza M K. Robust adaptive Kalman Filtering for systems with unknown step inputs and non-Gaussian measurement errors [J]. IEEE Transactions on Acoustics, 1986, 34(2):252-263
- [15] Xu J S, Qin Y Y, Peng R. Adaptive kalman filter fading factor selection method research [J]. Journal of Systems Engineering and Electronics, 2004, 26(11):1152-1154(in Chinese)
徐景硕,秦永元,彭蓉.自适应卡尔曼滤波器渐消因子选取方法研究[J].系统工程与电子技术,2004,26(11):1152-1154

(上接第 299 页)

根据类间区分度 D_R 的分布,确定了在噪声环境下表现更好的特征向量,并对这些向量进行融合,这种方法一定程度上忽略了干净语音下特征类间区分度,导致了融合声纹特征在干净语音环境下识别率反而下降的后果。

结束语 本文首先利用 Gammatone 滤波器组和 Meddis 模型组成的听觉外周模型改进了 MFCC 特征参数得到 GFCC,并研究了 PLPC 特征提取;然后为结合 GFCC 与 PLPC 在复杂噪声环境下表征说话人的优势,在不同信噪比噪声环境下统计两种特征类间区分度 D_R 分布,并在此基础上对 GFCC 和 PLPC 在特征域进行数据融合,得到 GFCC-PLPC 声纹特征;最后在不同信噪比环境下对 GFCC-PLPC 特征和单一特征进行实验。实验结果表明:GFCC-PLPC 融合特征参数比 MFCC 识别率有所提升,基于改进特征提取算法的说话人确认系统在噪声环境下具有比单一特征更强的鲁棒性,准确率得到有效提升,但在干净语音环境下存在误识别较严重的问题,需要寻求进一步解决的方法。

参 考 文 献

- [1] Pols L C W, Schouten M E H. Perception of tone, band, and formant sweeps[M]//The psychophysics of Speech Perception. Springer Netherlands, 1987;231-240
- [2] Hsu W, Sun J. The Effectiveness of Linear Prediction Residual to the Verification of Voiceprint and the Recognition of Chinese Tone[C]//IEEE International Symposium on Multimedia. IEEE, 2010;353-356
- [3] Patil H A, Basu T K. Comparison of subband cepstrum and Mel cepstrum for open set speaker classification[C]//Proceedings of the IEEE INDICON 2004 India Annual Conference, 2004. First. IEEE, 2004;35-40
- [4] Srinivasan A. Speaker identification and Verification using Vector quantization and Mel frequency Cepstral Coefficients[J]. Engineering and Technology, 2012, 4(1):33-40
- [5] Hermansky H. Perceptual linear predictive (PLP) analysis of speech[J]. The Journal of the Acoustical Society of America, 1990, 87(4):1738-1752
- [6] Li Y G, Ouyang X Z, Z F. A Study on Robust Speech Recognition Based on Gammatone feature[C]//National Conference on Man-Machine Speech Communication Chinese Information Processing Society of China, 2013;26-29(in Chinese)
李银国,欧阳希子,郑方.基于Gammatone特征的鲁棒语音识别研究[C]//第十二届全国人机语音通讯学术会议(NCMMSC'2013), 2013;26-29
- [7] Luo Y, Chen J, Zhang Y. An Auditory Feature Extraction Algorithm Based on Gama-chirp Filter Banks[J]. Information and Control, 2013, 42(5):590-594(in Chinese)
罗元,陈君,张毅.基于伽马啁啾滤波器组的听觉特征提取算法[J].信息与控制,2013,42(5):590-594
- [8] Makhoul J, Cosell L. LPCW: An LPC vocoder with linear predictive spectral warping[C]//IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '76). IEEE, 1976;466-469
- [9] Varga A, Steeneken H J M. Assessment for automatic speech recognition; II. NOISEX-92; A database and an experiment to study the effect of additive noise on speech recognition systems [J]. Speech Communication, 1993, 12(3):247-251
- [10] Bao H J, Zheng F. Combined GMM-UBM and SVM speaker identification system[J]. Tsinghua University (Science & Technology), 2007, 48(S1):693-698(in Chinese)
鲍焕军,郑方. GMM-UBM 和 SVM 说话人辨认系统及融合的分析[J].清华大学学报(自然科学版),2007,48(S1):693-698