

具有敏感区域的华南虎图像配准研究

王晓宇¹⁾ 陈吉红¹⁾ 李知践²⁾ 李曦¹⁾ 唐小琦¹⁾

¹⁾(华中科技大学国家数控系统工程技术研究中心, 武汉 430074) ²⁾(武汉理工大学自动化学院, 武汉 430070)

摘要 对数极坐标映射是一种具有刚体不变性的模板匹配算法,以它获得的刚体变换矩阵为初值经 LMA (Levenberg Marquardt algorithm) 迭代可以实现图像的高精度配准,但对于场景发生剧烈变化的图像对,上述方法无法获得正确的配准参数。为此,提出了一种具有敏感区域的图像配准算法,在 LMA 迭代中仅考虑落在人工选取的敏感区域中的像素点,通过对陕西林业厅公布的华南虎照片进行的图像配准实验表明了该算法的有效性。

关键词 对数极坐标映射 LMA 迭代 敏感区域

中图法分类号:TP391.41 文献标识码:A 文章编号:1006-8961(2008)11-2204-05

Image Registration of the South China Tiger with Sensitive Region Concerned

WANG Xiao-yu¹⁾, CHEN Ji-hong¹⁾, LI Zhi-jian²⁾, LI Xi¹⁾, TANG Xiao-qi¹⁾

¹⁾(National Numeric Control System Engineering Research Center, Huazhong University of Science and Technology, Wuhan 430074)

²⁾(Automation College, Wuhan University of Technology, Wuhan 430070)

Abstract The Log-polar mapping is a rigid invariant image matching algorithm, Levenberg Marquardt Algorithm (LMA) can calculate image registration parameters with high precision, but the LMA algorithm fails when facing an image pair with intense scene variance. In this paper we propose an image registration algorithm with sensitive region introduced in LMA phase, only pixels in the manually selected sensitive region are calculated in the iteration. The algorithm is used to register five representative photos of the Shanxi South China Tiger. The experiment shows that the method is a valuable improvement to the LMA algorithm.

Keywords Log-polar mapping, Levenberg Marquardt algorithm iteration, sensitive region

1 引言

2007年10月3日周正龙拍摄的陕西野生华南虎照片受到广泛关注和质疑,2007年11月9日《科学》杂志“随机样本”(random samples)栏目以《珍贵老虎照片沸沸扬扬》(RARE-TIGER PHOTO FLAP MAKE SFURFLY IN CHINA)为题,以“平面猫科动物?”(Flat cat?)为说明刊登了该照片。

图像配准的目的在于消除由于相机运动造成的空间变形,以突出图像中目标信息的变化。

图像配准算法包括基于特征匹配和基于灰度匹配两类,前者在预处理中提取的特征空间中进行搜索,后者一般直接在图像灰度信息中进行搜索和匹配。在基于灰度匹配的算法中,对数极坐标映射与 LMA(Levenberg Marquardt algorithm) 迭代结合的配准算法由于其具有刚体不变性、匹配能力强、配准参数准确而受到广泛关注^[1-5]。

然而对于场景发生剧烈变化的图像对,直接运用上述方法无法获得正确的配准参数。为此提出了一种具有敏感区域的 LMA 迭代算法,在迭代中仅考虑落在人工选取的敏感区域中的像素点,虎照配准

实验验证了该算法的有效性。

2 对数极坐标映射与 LMA 迭代

该算法包括两个阶段,第 1 阶段是在参考图像和观察图像上按照对数极坐标映射进行模板匹配,获得刚体变换矩阵;第 2 阶段执行 LMA 迭代,以上述刚体矩阵为初值求解平面投影变换矩阵,最终获得亚像素精度的配准参数。

2.1 对数极坐标映射与相位相关法

对图像中笛卡儿坐标为 (x, y) 的像素点,在复平面对应于点 $z = x + yi = \rho e^{i\theta}$,其对数极坐标为 $\omega = \log(\bar{z})$,对任意的 $z \neq 0$,定义 $\text{Re}(\omega) = \log(\rho)$, $\text{Im}(\omega) = 2k\pi + \theta$ 为点 (x, y) 的对数极坐标映射,其中, $\theta \in [0, 2\pi)$ 。

定义表明如果 z 空间物体相对于坐标原点发生了刚体变化,例如放大 ρ_0 倍,旋转 θ_0 角度,则变化后相应新的极坐标为 $(\rho_0\rho, \theta + \theta_0)$,取对数极坐标映射后为

$$\begin{cases} u(\rho, \theta) = \log \rho_0 + \log \rho \\ v(\rho, \theta) = \theta + \theta_0 \end{cases} \quad (1)$$

也就是说 z 空间中物体的比例变化相当于对数极坐标 w 空间中物体在水平方向发生位移,而旋转变化则相当于垂直方向发生位移,并以 2π 为周期。即当变换中心与目标质心重合时,对数极坐标映射具有刚体不变性,这是对数极坐标映射的最重要性质^[1]。

傅里叶平移定理表明,如果图像 $f_2(x, y)$ 是图像 $f_1(x, y)$ 经平移 (x_0, y_0) 后的图像,即 $f_2(x, y) = f_1(x - x_0, y - y_0)$,设 F_1, F_2 分别为图像 f_1, f_2 的傅里叶变换,则频域中两个图像的互能量谱:

$$\frac{F_1(\xi, \eta) F_2^*}{\|F_1(\xi, \eta) F_2^*\|} = e^{2i\pi(\xi x_0 + \eta y_0)}$$

的反变换是一个在偏移位置的冲击函数 $\sigma(x - x_0, y - y_0)$,其中 F_2^* 是 F_2 的复共轭。

图像的粗匹配过程是:首先在参考图像 f_1 中以 (x, y) 为中心获得一圆形参考子图 \tilde{f}_1 ,将其通过对数极坐标映射到 w 空间,得到 \tilde{f}_1^{lp} (上标 lp 表示对数极坐标映射),对待匹配观察图像 f_2 上所有可能的坐标点 (u, v) ,以改点为中心取与参考子图相等大小的观察子图并映射到对数极坐标得到 \tilde{f}_2^{lp} 。

运用相位相关法计算观察子图与参考子图的偏

移 $(\Delta x, \Delta y)$ 及峰值大小,并记录到一个 2 维结构数组中,具有最大峰值的点 (u, v) 即为观察子图中心, (u, v) 与 (x, y) 的位移为平移分量;观察子图与参考子图的峰值偏移 $(\Delta x, \Delta y)$ 分别对应缩放和旋转分量:

$$s = \begin{cases} Base^{\Delta x \times r_resolution} & \Delta x < \frac{D}{2} \\ 1 & \Delta x > \frac{D}{2} \\ Base^{\Delta x \times r_resolution} & \Delta x < \frac{D}{2} \end{cases} \quad (2)$$

$$\theta = \Delta y \times \theta_resolution$$

式中, $Base$ 为对数的基, D 为模板大小, $r_resolution$ 和 $\theta_resolution$ 分别是对数极坐标子图中两方向的分辨率。需要注意,由于旋转中心是 (u, v) ,因此需要执行文献[6]所示的复合变换得到刚体变换矩阵。

2.2 LMA 迭代

平面投影变换满足以下关系:

$$\begin{cases} x'_i = \frac{m_0 x_i + m_1 y_i + m_2}{m_6 x_i + m_7 y_i + 1} \\ y'_i = \frac{m_3 y_i + m_4 x_i + m_5}{m_6 x_i + m_7 y_i + 1} \end{cases} \quad (3)$$

式中, m_0, \dots, m_7 为投影变换参数, LMA 迭代的目标函数为使得两幅图像重叠区域的累加平方差(SSD)最小:

$$\chi^2(\alpha) = \sum_{i=1}^N [f_1(x_i, y_i) - f_2(x'_i, y'_i)]^2 \quad (4)$$

LMA 求解的是以刚体变换矩阵为初值,确定式中的 m_i 实现参考图与样本图的配准,并使得 SSD 达到最小的非线性最小平方问题,是一种有阻尼的高斯牛顿法,其迭代方向 h_{lm} 由下式得到:

$$(J^T J + \mu I) h_{\text{lm}} = -g \quad (5)$$

式中, $\mu > 0$ 是阻尼因子, J 为 Jacobian 矩阵, $J^T J = H$ 为 Hessian 矩阵, g 为梯度矢量。

LMA 方法综合了梯度法和高斯牛顿法的优点,在搜索过程中通过修改阻尼因子 μ 来动态决定搜索方向,有效地克服了梯度法和高斯牛顿法只能适应特定区域初值的缺陷。

式(5)中 Hessian 矩阵 $H = [a_{kl}]$ 和梯度向量 $g = [g_k]$ 计算如下:

$$\begin{aligned} \frac{\partial e_i}{\partial m_0} &= \frac{x_i}{D_i} \frac{\partial f_2}{\partial x'} & \frac{\partial e_i}{\partial m_1} &= \frac{y_i}{D_i} \frac{\partial f_2}{\partial x'} & \frac{\partial e_i}{\partial m_2} &= \frac{1}{D_i} \frac{\partial f_2}{\partial x'} \\ \frac{\partial e_i}{\partial m_3} &= \frac{x_i}{D_i} \frac{\partial f_2}{\partial y'} & \frac{\partial e_i}{\partial m_4} &= \frac{y_i}{D_i} \frac{\partial f_2}{\partial y'} & \frac{\partial e_i}{\partial m_5} &= \frac{1}{D_i} \frac{\partial f_2}{\partial y'} \\ \frac{\partial e_i}{\partial m_6} &= -\frac{x_i}{D_i} \left(x'_i \frac{\partial f_2}{\partial x'} + y'_i \frac{\partial f_2}{\partial y'} \right) \end{aligned} \quad (6)$$

$$\frac{\partial e_i}{\partial m_7} = -\frac{x_i}{D_i} \left(x_i' \frac{\partial f_2}{\partial x'} + y_i' \frac{\partial f_2}{\partial y'} \right)$$

式中, $D_i = m_6 x_i + m_7 y_i + 1$, $a_{kl} = \sum_i \frac{\partial e_i}{\partial m_k} \frac{\partial e_i}{\partial m_l}$, $g_k = -2 \sum_i e_i \frac{\partial e_i}{\partial m_k}$, 带入式(5)即可得到 Δm 。

2.3 具有敏感区域 LMA 迭代

在陕西林业厅发布的 40 张照片^[7]中选择编号为 06, 24, 27, 29, 40 共 5 张有代表性的图像进行配准实验, 其中编号 06(图 1(a))为林业厅作为新闻发布的效果最好的照片, 将该照片作为参考图像与其他 4 幅进行配准, 并将编号为 24 的无遮挡图像与年画虎照片进行配准。配准后将参考图像 f_1 、配准后的观察图像 f_2' 及两者的拼接图像(灰度值为两幅图像对应点灰度值各取 50%) 构成 GIF 动画, 以观察老虎姿态变化。

在对编号为 27(图 1(b))图像进行拼接时, 直接运用对数极坐标映射与 LMA 迭代算法无法得到的正确的配准参数, 生成动画后可用肉眼观察到老虎存在 2 个像素左右的微小移动, 而且这个差异不会随迭代次数的增加而消除。



(a)参考图像(06)

(b)待拼接图像(27)



(c) 06与27配准图像

图 1 敏感区域与配准结果

Fig. 1 Registration with sensitive region

样本图像具有相同内容的重叠区域的情况, 而且重叠区域中的每个点均可按投影参数实现一一对应。但在拼接图像图 1(c)中, 两幅图像的重叠区域中只有不到 10% 面积的像素点指向真正相同的目标(老虎), 而其他 90% 以上的像素点指向老虎周围的环境, 而从图中可以看到两张照片中老虎所处的环境完全不同, 两处环境中杂乱的树叶根本不存在着对应关系, 因此如果直接采用前述的图像拼接算法, 占面积 90% 以上的不可匹配像素参与到 LMA 迭代中, 因此算法得到错误的配准参数。

为此在参考图像中定义与作为配准目标的敏感区域, 如图 1(a)所示, 图中敏感区域是两个多边形的并集, 在算法精匹配阶段, 对式(4)中的像素对 $(x_i, y_i) \in f_1$ 和 $(x_i', y_i') \in f_2$, 如果 (x_i, y_i) 落在敏感区域内, 则该对像素参与 Hessian 矩阵和梯度向量的计算(式(6)), 否则忽略该像素对的贡献。

加入敏感区域限制后, 得到的拼接图像如图 1(c)所示, 可以根据图像清晰度进行配准效果的判别, 从图中可以看到, 虎身上的特征(包括在花纹、口鼻等器官)清晰、位置一致, 不存在由空间差异导致的虚影, 而树叶及虎身上被树叶遮挡的特征, 则存在明显的虚影, 说明在该区域不能实现配准; 拼接图像与原图像构成的动画不存在肉眼可辨的空间变形, 两幅图像中虎照可以实现精确配准。

3 虎照配准实验及误差分析

加上敏感区域限制后, 对上述编号的所有图像对均可以实现准确配准, 老虎身上特征均无虚影(图 2); 原图及拼接图构成的动画中, 老虎身上的花纹等特征均不存在肉眼可辨的位置移动。

图 3 是视差估算示意图, 设老虎深度 AG 为 1m (即虎脸与后腿相距 1m), 拍摄距离 GB 为 20m, 两次拍摄相机间距 BC 为 0.5m, 佳能 EOS 400D 尺寸为 22.2×14.8 mm 的 CCD 上, 如果 20m 外 1m 的目标在 CCD 上占据约 1.36mm (按照 3888×2592 分辨率 170 像素计算), 可得到像距 BD 约为 27mm, 图 3 所示的视差 EF 在 5 个像素以上(约为 0.032mm)。而且上述数值均为保守的数值, 按照拍摄场景重建新闻中的数据是拍摄距离为 15.5m, 相机位移在 1m 以上, 因此实际视差对应的偏移像素数更大。

分析后发现, 前述算法只适用于参考图像和



图 2 拼接图

Fig. 2 Mosaic images

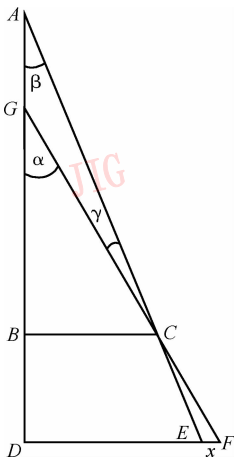


图 3 视差估算示意图

Fig. 3 Parallax estimation sketch

上述分析说明如果拍摄目标为一 3 维华南虎,按照新闻中提供的拍摄距离、相机 CCD 尺寸及分辨

率,估算的 0.5m 位移并按照老虎头部配准后,在虎后腿处,1m 的深度差在两次拍摄中形成的视差至少为 5 个像素。

图 1(c)及图 2 所示的拼接图像不存在肉眼可辨的特征移动,为证实这一点,将编号 06 与 27(图 1(a)、(b))虎照的后腿处局部敏感区域的细节绘制成图 4,图 4 中后腿花纹特征下部(图 4(a)、(b)中两点位置及图 4(c)、(d)中直线位置)由于不存在遮挡,是可匹配的特征,由图 4(a)、(b)可见配准误差小于 1 像素;而花纹特征上部由于在 27 号图中出现了遮挡的植物,因此是不可匹配的特征,无法估算配准误差。

文献[5]指出对无视差图像经 LMA 迭代可获得亚像素精度的配准参数,在增加敏感区域限制后这一点在实验中得以实现,图 2 所示的其他图像对配准误差均小于 1 个像素。

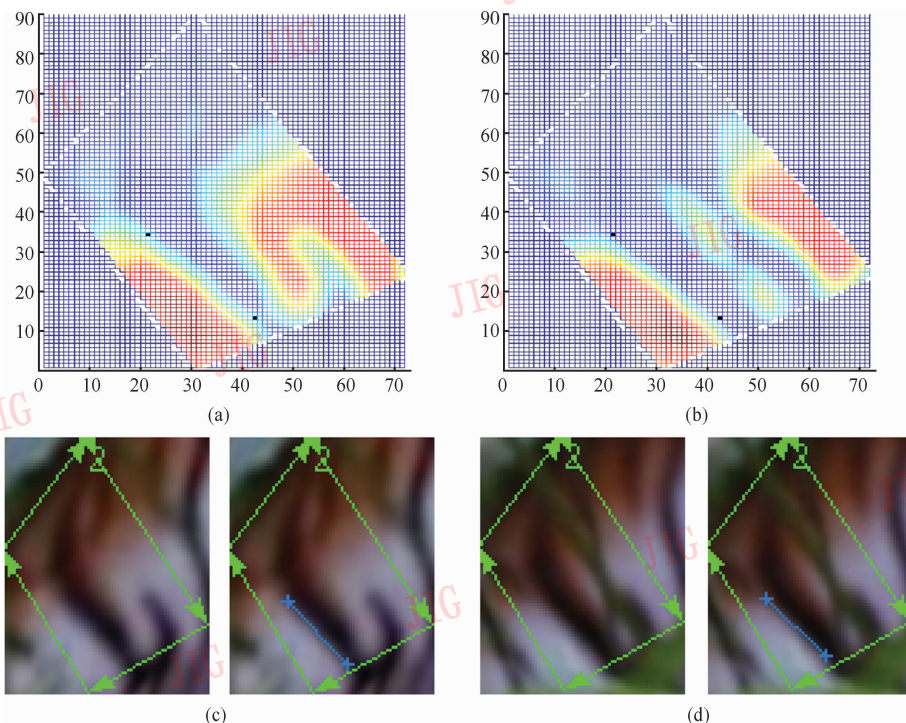


图 4 后腿部分特征细节

Fig. 4 Local details of features on back leg

4 结 论

通过上述配准实验可以得到以下结论:

(1) 图像配准可以消除由于相机运动产生的图像变形并突出老虎的姿态变化, 结果显示老虎的姿态改变小于 1 像素;

(2) 保守估计下虎头与后腿间的深度差应形成 5 个像素的视差, 但配准结果显示特征位置误差小于 1 个像素;

(3) 虎照均可由平面投影变换模型准确配准, 目标无可测得的深度信息。

实验结果表明本文配准算法可以解决场景发生剧烈变化情况下的图像配准, 是对数极坐标映射与 LMA 迭代算法的有益发展。

参考文献 (References)

- Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation, and scale invariant image registration [J]. IEEE Transactions on Image Processing, 1996, 5(8): 1266 ~ 1271.
- George Wolberg, Siavash Zokai. Robust image registration using log-

polar transform [J]. IEEE Transactions on Image Processing, 2000, 6(4): 493 ~ 496.

- Chen Chao-yang, Ding Ming-yue, Zhou Cheng-ping. A target tracking method based on log-polar coordinate mapping [J]. Chinese Journal of Electronics, 2001, 29(9), 1173 ~ 1175. [陈朝阳, 丁明跃, 周成平. 基于对数极坐标映射的目标跟踪方法 [J]. 电子学报, 2001, 29(9), 1173 ~ 1175.]
- Li Zhong-xin, Mao Yao-bin, Wang Zhi-quan. A method of image mosaicing using log-polar coordinate mapping [J]. Journal of Image and Graphics, 2005, 10(1): 59 ~ 63. [李忠新, 茅耀斌, 王执铨. 基于对数极坐标映射的图像拼接方法 [J]. 中国图象图形学报, 2005, 10(1): 59 ~ 63.]
- George Wolberg, Siavash Zokai. Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations [J]. IEEE Transactions on Image Processing, 2005, 14(4): 1422 ~ 1434.
- Sun Jia-guang. Computer Graphics (3rd Edition) [M]. Beijing: Tsinghua University Press, 2000: 369. [孙家广著. 计算机图形学 (第 2 版) [M]. 北京: 清华大学出版社, 2000: 369.]
- Wang Feng. NetEase Open All Tiger Photos [EB/OL]. <http://news.163.com/special/00012FD0/tigerpho071126.html>, 2007-11-26. [王岚. 网易独家披露全套华南虎数码照片 [EB/OL]. <http://news.163.com/special/00012FD0/tigerpho071126.html>, 2007-11-26.]