

一种基于对数极坐标变换的快速目标识别算法

严江江,丁明跃,周成平

(华中科技大学 图像信息处理与智能控制教育部重点实验室,湖北 武汉 430074)

摘要:在目标识别的过程中,观察目标图像相对于基准目标图像会存在尺度、方向和位置的变化,使得识别速度和准确率降低.针对这一问题,提出了一种快速对数极坐标变换算法,加快了从笛卡尔坐标转换到对数极坐标的过程.通过采用双轴投影相似度分析算法对目标图像进行匹配,进一步加快了识别速度,同时保证了匹配的可靠性.理论分析和试验结果表明,该算法在计算效率和目标识别正确率方面具有较好的性能.

关键词:对数极坐标变换;目标识别;图像匹配;快速变换

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1673-4785(2008)04-0370-07

A fast target recognition algorithm based on LPT

YAN Jiang-jiang, DING Ming-yue, ZHOU Cheng-ping

(State Education Commission Key Lab for Image Processing and Intelligent Control, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: In the course of target recognition, recognition speed and accuracy decreases if the scale, orientation and position of the target image changes with respect to the reference image. This paper presents a fast log-polar transformation (LPT) algorithm, which speeds up processing from Cartesian coordinates to log-polar coordinates as well as speeding up image matching by using a dual-axis projection algorithm. The new algorithm greatly accelerates target recognition and ensures the reliability of matches. Theoretical analysis and experiments showed that this method has better performance in both recognition speed and accuracy.

Keywords: log-polar transformation; target recognition; image matching; fast transformation

在图像处理领域,目标识别是一项极其重要的研究内容,其应用包括机器视觉、成像制导和文字识别等.一般的目标识别过程是计算观察图像与基准图像的相似度,这种方法称之为“模板匹配”.通常情况下目标图像相对于基准目标图像会发生尺度、方向、位置的变化,通过对数极坐标变换(log-polar

transformation, LPT)可以解决这个问题.对数极坐标变换源于人类视觉中视网膜皮层映射关系的研究^[1-2],人们将其广泛应用于目标识别、目标跟踪、运动检测、脸部检测等领域^[3-5].笛卡尔坐标系下目标的尺度和旋转变换映射到对数极坐标系,表现为目标区域的水平和垂直方向的平移,而目标的轮廓区域保持不变.利用这种特性可以对目标进行匹配和识别.

收稿日期:2007-12-19.

基金项目:国家自然科学基金重点资助项目(60135020FF030405).

通信作者:严江江. E-mail: yjsalriver@yahoo.com.cn

通常目标识别算法在速度上都有很高的要求,本文提出了一种新的快速 LPT 算法,通过减少采样点数目和采用查找表方法来减少运算量.在此基础上进一步提出了一种双轴投影相似度分析算法对目标进行匹配识别,通过统计轴向投影量将二维图像的匹配识别转换为一维的轴向投影匹配识别,可同时获得较高的计算效率和较好的正确识别率.

1 快速识别算法及其实现

笛卡儿坐标系中的点 (x, y) 映射到对数极坐标系中的点 (r) 的计算公式为



图 1 原始目标图像及其 LP 图像以及构造 LP 图像的原图像采样点图

Fig 1 Input image, log-polar image, and sampling points of input image to construct the log-polar image

本文只针对二值图像进行处理.从图 1(c)中看到,极坐标原点处的采样最为密集,采样密度沿中心向外围递减,这样 LP 图像对目标的平移将会相当敏感.将采样原点取在目标形心的位置来消除目标平移的影响,物体形心的确定参见文献 [7].图 1(b)中黑白相间的地方就是图 1(a)中目标轮廓区域在 LP 平面的映射.在 LP 图像中,往往只对目标轮廓感兴趣,因为它反映了目标的形状,而对靠近采样原点的图像部分则不感兴趣.如果对靠近采样原点的区域不采样或者少采样,就可以减少 LPT 的运算量.因此,只对满足式 (3)的点作式 (2)的变换.

$$r = \dots \tag{3}$$

以 R 代表 LP 图距离轴长度, M 表示原图像中离采样原点最远的距离, S 表示目标边界离采样原点最近的距离,则 通过式 (4)确定:

$$= \frac{\ln(S)}{\ln(M)} \cdot R \tag{4}$$

令 $S = M \cdot k$, 其中 $0 < k < 1$, 则式 (4)可写成:

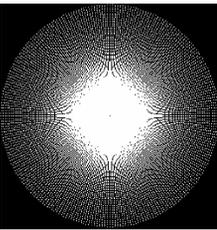
$$r = \ln(\sqrt{x^2 + y^2}), \quad \theta = \arctan(y/x) \tag{1}$$

由于图像是离散栅格化表示的,原图像的像素直接生成对数极坐标 (LP) 图像的输出像素,这无法保证 LP 图像的像素都能被映射到.本文采用间接重采样法 [6] 来获得 LP 图像.

首先设定 LP 图像的大小,然后计算 LP 图像中的点 (r) 映射到原图像中的点坐标 (x, y) ,其计算公式为

$$x = \exp(r) \cdot \cos(\theta), \quad y = \exp(r) \cdot \sin(\theta) \tag{2}$$

最后将 (x, y) 的值取整后赋给 (r) .



(c) 原图像采样点图

$$= \dots \cdot R \tag{5}$$

如果 $k = 0.5$, 有 $S = M^{0.5} = \sqrt{M}$, 这时 $k = R/2$, 即 LPT 的运算量减少了一半.实际应用中, S 需要计算才能得到,但基准图的 S 是预先知道的,可以通过基准图的 S 估算出 k 的值.考虑到图像是栅格化表示, (r) 为 LP 图中整数值的坐标,而 LP 图像的大小是事先确定的,则每次计算的输入参数和结果也是确定的.如果事先使用一个表来记录式 (2) 中的输入参数和结果,在以后计算的过程中就只需要通过输入参数来查找表中对应的结果即可,这样式 (2) 中复杂的自然幂函数和三角函数运算就转换成了读取内存地址中内容的操作 (相当于加法运算),这将大大加快获得 LP 图像的速度.

由于极坐标原点取在目标的形心位置,消除了目标平移的影响,因此只需要考虑目标发生尺度和旋转变化的情况.极坐标系下的点 (r, θ) 转换到对数极坐标下的对应点为 (r) , 即 $r = \ln(\dots)$. 设目标的尺度变化因子为 k , 旋转变因子为 θ , 则 (r) 发

生尺度和旋转变化后的对应点为 (r, θ) , 可由式

(6) 计算得到:

$$r = \ln(k \cdot r_0) = \ln(k) + \ln(r_0) = \ln(k) + r_0$$

$$\theta = \theta_0 + \alpha$$
(6)

可以看出, 目标的尺度变化相当于在距离轴上平移了 $\ln(k)$ 个单位, 而目标的旋转变化相当于在角度轴上平移了一个单位, 图 2 清晰地表明了这种关系.

目标的识别通过 LPT 转化成 LP 图与基准 LP 图的目标轮廓的匹配问题. 如果在一个场景中存在多个待识别目标, 则分别对每一个目标进行匹配识别即可. 本文基准图为图 1(a), 基准 LP 图为图 1(b), LP 平面水平方向为距离轴方向, 垂直方向为角度轴方向. 为了减少目标轮廓提取的计算量, 本文提出了一种逆向距离轴投影分割方法, 用 $f(r, \theta)$ 和 $L(r, \theta)$ 分别表示 LP 图在距离轴和角度轴上的投影统计量.

$$f(r, \theta) = \sum_{i=1}^N L(r, \theta_i)$$

$$L(r, \theta) = \begin{cases} 0, & f(r, \theta) = 0; \\ 1, & f(r, \theta) > 0 \end{cases}$$
(7)

式中: $L(r, \theta)$ 为点 (r, θ) 的值. 在 LP 图中, 目标轮廓右边距离轴投影统计量为 0, 而目标轮廓左边距离轴投影统计量为 N , 从距离轴最大值逆向开始计算距离轴投影, 通过 0 和 N 这 2 个阈值就可以分割出目标轮廓区域, 分割完毕即停止计算. 如图 3 所示, 图 3 (a) ~ (c) 上半部分为距离轴投影曲线, 下半部分为通过计算投影统计量分割出来的目标轮廓区域.

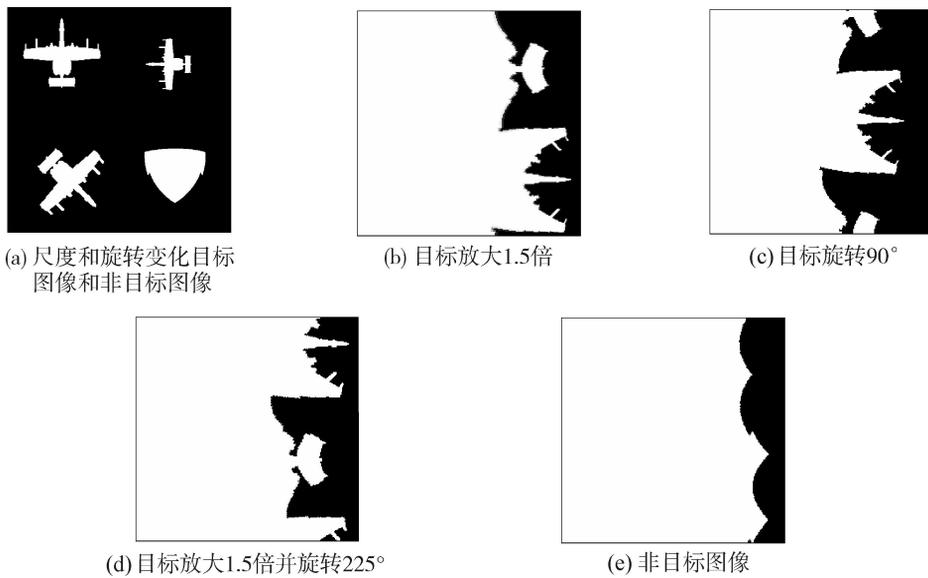


图 2 尺度和旋转变化对应的对数极坐标变换图

Fig 2 The LPT image for object's scaling and rotating change

用经典相关匹配算法^[8]是不能直接对分割出来的目标轮廓区域进行比较的, 需要作一些相应处理. 式 (2) 计算时存在坐标取整过程, 目标尺度和旋转变化可能会造成目标轮廓区域的采样误差, 造成分割出的宽度不一致, 如果宽度相差过大, 则直接判定不是目标. 如果满足:

$$|W_p - W_q| \leq \epsilon$$
(8)

则作下一步的相似度计算. 其中, W_p 和 W_q 分别为 LP 图和基准 LP 图的目标轮廓区域宽度, ϵ 为先验知识, 本文中 ϵ 取值为 3. 为了对目标轮廓区域进行匹配分析并减少计算量, 设计了双轴投影相似度分析算法, 首先用式 (7) 得到目标轮廓区域在距离轴和角度轴上的投影, 再与基准 LP 图相应投影进行匹配分析. 文献 [9] 对距离轴投影进行相似度分析

就得出目标匹配的结果,虽然保证了运算量的降低,但不具有鲁棒性,因为仅仅由距离轴投影相似并不能完全说明图像的相似.本文首先判断距离轴投影,

在给定的阈值下不相似,则直接判定不是目标,相似则再比较角度轴投影.角度轴投影如图 4 所示.

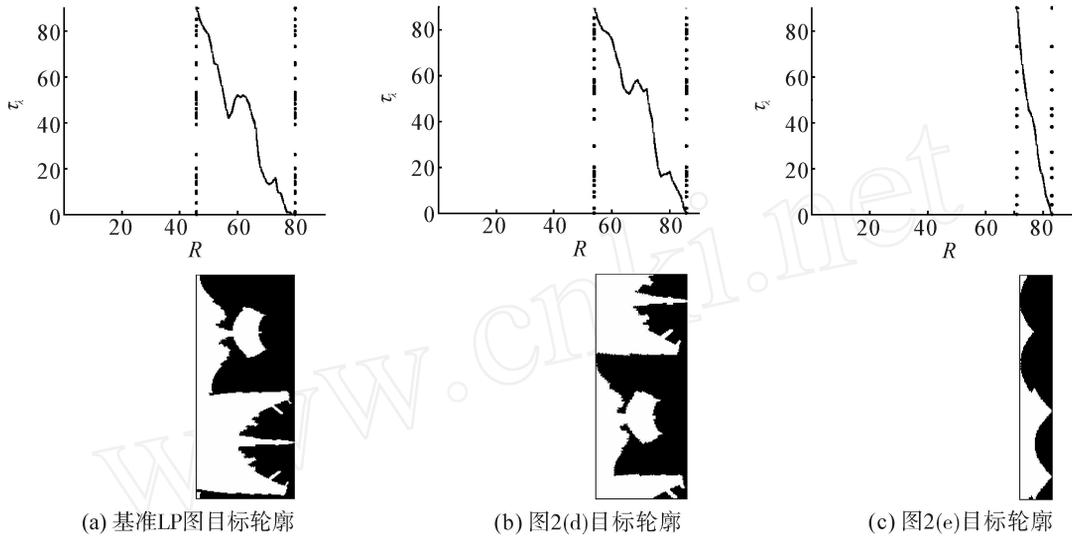


图 3 逆向距离轴投影分割方法

Fig 3 The segmentation method based on converse distance axis projection

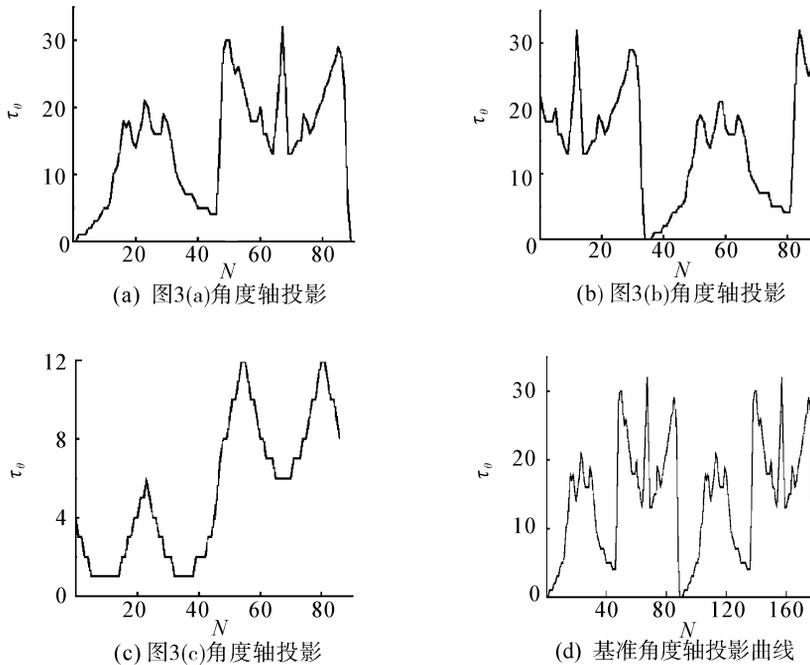


图 4 角度轴投影曲线

Fig 4 The projective curve of the angle axis

将图 4(a)进行复制扩展,就得到了基准角度轴匹配曲线,如图 4(d)所示.在给定的阈值下,距离轴和角度轴投影与基准投影均相似,则判定是目标;否则不是.相似度比较算法采用常用的平均绝对差算法,公式定义为

$$D_l = \frac{1}{m} \cdot \sum_{i=1}^{l+m+1} |N_p(i) - N_q(i)| \quad (9)$$

式中:如果是距离轴投影比较, $m = \min(W_p, W_q)$, $l = 1, 2, \dots, |W_p, W_q| + 1$;如果是角度轴投影比较, $m = W_p$, $l = 1, 2, \dots, W_p$. $N_p(i)$ 、 $N_q(i)$ 分别表示 LP 图和基准 LP 图的轴向投影量,设定相似度阈值为 τ ,如果距离轴和角度轴投影均有 $\min(D_l) \leq \tau$,则停止计算,并认定是目标.

2 算法复杂度与分析

要得出识别结果,需 4 个步骤:1)得到观察图像 LP 图;2)分割出观察图像 LP 图中的目标轮廓区域;3)计算目标轮廓区域在距离轴和角度轴上的投影统计量;4)与基准 LP 图作比较.

分别对每步所需计算量进行分析,1)查找表中取内存地址内容运算相当于加法运算,需要 $(1 -) \cdot R \cdot P$ 次加法运算;2)计算轴向投影量相当于加法运算,最多需要 $(1 -) \cdot R \cdot P$ 次加法运算;3)距离轴投影在第 2 步中已经计算,这一步需要 $W_p \cdot P$ 次加法运算;4)由式 (9)取极限情况,对距离轴投影,需要 $2 \cdot (|W_p - W_q| + 1) \cdot \min(W_p, W_q)$ 次加法运算和 $(|W_p - W_q|)$ 次乘法运算,对角度轴,需要 $(2 \cdot P) \cdot P$ 次加法运算和 P 次乘法运算;将加法和乘法都作为基本运算,则本文算法共需要 $[2(1 -)R + 2P + W_p + 1] \cdot P + (|W_p + W_q| + 1) \cdot [2\min(W_p, W_q) + 1]$ 次基本运算.

3 试验结果与分析

常用 LPT 算法是自然幂和三角函数运算组成的复杂运算,将自然幂函数、正弦函数和余弦函数展开成泰勒级数,分别记为 N_E 、 N_S 、 N_C . 根据泰勒展开

式,一次 N_E 运算相当于 $k + \sum_{n=2}^k (2n + 1)$ 次基本运

算,一次 N_S 运算相当于 $k + \sum_{n=2}^k (2(2n - 1) + 1)$ 次基本运算,一次 N_C 运算相当于 $k + \sum_{n=2}^k [2(2n) + 1]$ 次基本运算,其中 k 为计算机中泰勒级数计算到足够精确时 n 的取值,则常用 LPT 算法共需要 $4k + (12n + 2)$ 次基本运算. 观察图像大小为 256×256 , LP 图像 90×90 , 取 $W_p = W_q = 30$, $\tau = 0.5$, $R = P = 90$, k 取值为 15. 表 1 为算法的计算效率对比情况. 在以上试验条件下,本文 LPT 算法相对于常用 LPT 算法速度提高了大约 3 000 倍,匹配算法相对于常用相关匹配算法速度提高了 30 倍左右.

表 1 算法的计算效率对比

Table 1 The comparison of calculational efficiency

算法	基本运算次数
常用 LPT 算法	12 279 600
本文 LPT 算法	4 050
相关匹配识别算法	486 090
本文匹配算法	16 351

对于有不同噪声背景的观察图像,生成 LP 图像时会将噪声带入,如图 5(a) ~ (b),采用填充和去除孤立点的方法可消除噪声的影响,如图 5(c)所示.

用图 2(b) ~ (d)、去除噪声后的图 5(c)与基准 LP 图作相似度计算,结果如表 2 所示. 其中, $W_q = 32$ W_p 为观察 LP 图目标轮廓区域宽度,距离轴和角度轴的 $\min(D_l)$ 为式 (9) 计算出来的各自平均绝对差的最小值.

表 2 相似度分析

Table 2 The analysis of similarity

图序	W_p	距离轴 D_l	角度轴 D_l
图 2(b)	31	1.903 2	0.822 2
图 2(c)	33	3.750 0	2.355 6
图 2(d)	32	2.218 8	1.344 4
图 5(c)	31	3.300 0	2.677 8

本文共采用 3 组不同图像进行了实验. 第 1 组是 40 张目标作各种尺度和旋转变化的图像,使用最近邻插值方法;第 2 组是 90 张加入了各种不同噪声

的目标变化图,加入的噪声包括高斯噪声、椒盐噪声和加性噪声;第 3 组是 50 张非基准目标图像,其中 3 张有代表性图像如图 6 所示;实验结果见表 3,每

一组中正确识别的图像数除以每组总图像数即为识别正确率.

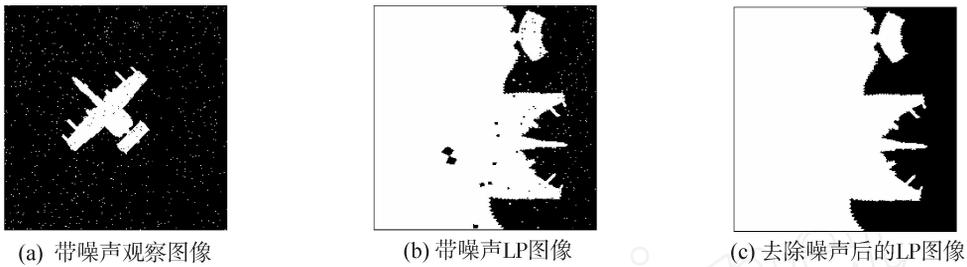


图 5 有噪声的观察图像及其带噪声 LP 图像和去除噪声后的 LP 图像

Fig 5 The input image, the log-polar image with noise and the log-polar image with non-noise

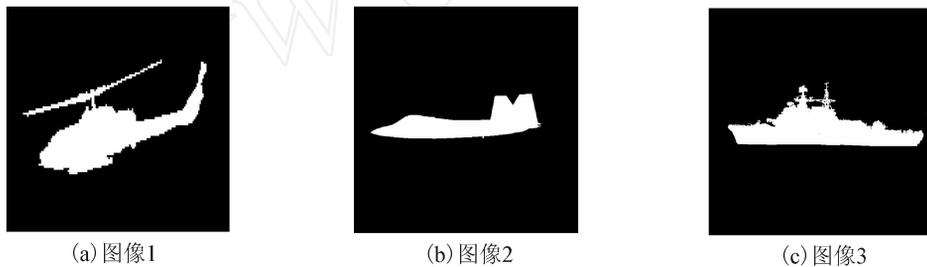


图 6 非基准目标图

Fig 6 Images of non-object

表 3 目标识别正确率

Table 3 Results of the recognition correct rate

组别	正确识别率 /%			
	$r = 2$	$r = 3$	$r = 4$	$r = 5$
第 1 组	32.5	75	95	97.5
第 2 组	28.8	75.5	94.4	97.7
第 3 组	100	100	100	100

从表 3 可以看到,噪声对识别正确率几乎没有影响.因为本文方法是针对目标轮廓的,在噪声不破坏目标轮廓提取的情况下是不影响识别正确率的.阈值 r 的选取很重要,在试验中取 4 或 5 就可以达到较高的识别正确率.对于非基准目标图,大多数不满足式 (8),少量满足式 (8) 的相似度又不满足阈值条件,所以取得了很高的识别正确率.

4 结束语

针对常用的对数极坐标变换算法,本文提出了

一种快速的改进算法,使其在计算效率上相对于常用算法有了较大的提高.同时本文利用对数极坐标平面目标的尺度和旋转不变性,提出了一种双轴投影相似度分析算法对分割出来的目标轮廓区域进行相似度分析,增加了其可靠性.将二维图像的匹配运算转换成 2 个轴向的一维投影匹配运算,减少了图像匹配的运算量,对目标识别尤其是快速目标识别提供了一个有效的途径.

参考文献:

[1] SCHWARTZ E L. Spatial mapping in the primate sensory projection: analytic structure and relevance to perception [J]. Biological Cybernetics, 1977, 25: 181-194.

[2] SCHWARTZ E L, GREVE D N, BONMASSAR G. Space-variant active vision: definition, overview and example [J]. Neural Networks, 1995, 8 (7): 1297-1308.

[3] BISHAY M, PETERS R A I, KAMWAMURA K. Object detection in indoor scenes using log-polar mapping [C] //

Proceedings of IEEE International Conference on Robotics and Automation San Diego, CA, USA, 1994: 775-780.

[4] SILVA C, SANTOS-VICTOR J. Egomotion estimation using log-polar images [C]// Sixth International Conference on Computer Vision Bombay, India, 1998: 967-972.

[5] ESCOBAR M J, RUIZ-DEL-SOLAR J. Biologically based face recognition using Gabor filters and log-polar images [C]// International Joint Conference on Neural Networks (IJCNN2002). Honolulu, USA, 2002: 1143 - 1147.

[6] HOTTA K, KURITA T, MISHIMA T. Scale invariant face detection method using higher-order local autocorrelation features extracted from log-polar image [C]// Proceedings of the 3rd IEEE International Conference on Face & Gesture Recognition Nara, Japan, 1998: 70-75.

[7] 冯月霞, 刘波. 用 TMS320C50 实现图像目标的形心捕获 [J]. 光子学报, 2001 (4): 455-459.

FENG Yuexia, LU Ba. The poird capturing of image object is realized using TMS320C50 [J]. Acta Photonica Sinica, 2001 (4): 455-459.

[8] TANG Minshou. The application of correlation matching technique in image guidance [C]// Proceeding of CSAA 1st Electronic Conference of the Youth Los Angeles, USA, 1994.

[9] 王立, 李言俊, 张科. 对数极坐标变换识别算法在

成像制导中的应用 [J]. 宇航学报, 2005, 26 (3): 330-333.

WANG Li, LI Yanjun, ZHANG Ke. The application of target recognition algorithm based on log-polar transformation for imaging guidance [J]. Journal of Astronautics, 2005, 26 (3): 330-333.

作者简介:



严江江,男,1979年生,博士研究生,主要研究方向为飞行器任务规划、智能搜索、计算机视觉等.发表学术论文 6 篇.



丁明跃,男,1961年生,特聘教授、博士生导师,IEEE高级会员,SPIE会员,湖北省青年科协副会长,湖北省宇航学会副理事长,主要研究方向为图像处理与分析、计算机视觉及应用、医学图像处理和任务规划等.发表学术论文 200 余篇,其中被 SCI, EI, ISTP 收录 130 余篇.



周成平,男,1957年生,副教授,主要研究方向为图像处理、计算机视觉、任务规划等.先后获航天部、教育部科技进步三等奖各一项,发明专利一项;发表学术论文 30 余篇.

2008 International Conference on Intelligent Computation Technology and Automation

2008年智能计算技术与自动化国际会议

2008 International Conference on Intelligent Computation Technology and Automation (ICICTA 2008) will be held on 20 ~ 22 October, 2008 in Changsha, China. ICICTA08 aims to provide a high-level international forum for scientists, engineers, and educators to present the state of the art of intelligent computation and automation research and applications in diverse fields. The conference will feature plenary speeches given by renowned scholars and regular sessions with broad coverage.

The conference proceedings will be published by IEEE Computer Society. All papers accepted will be included in IEEE Xplore and arranged for indexing by both EI Compendex and ISTP.

Email: icicta@126.com

Website: <http://www.icicta.org/index.asp>