

文章编号 1004-924X(2009)09-2276-07

# 图像特征点提取及匹配技术

杨晓敏, 吴 炜, 卿 粼波, 华 骅, 何小海

(四川大学 电子信息学院 图像信息研究所, 四川 成都 610064)

**摘要:** 针对图像特征提取与匹配的适应性和准确性问题, 将尺度不变特征变换算法(SIFT)应用到图像的特征点提取与匹配中。SIFT算法可在尺度空间寻找极值点, 提取对图像尺度和旋转变化具有不变性, 对光照变化和图像变形具有较强适应性的特征点及其特征描述。首先, 采用SIFT算法提取图像的特征点及其描述, 然后, 采用基于置信度的匹配算法进行特征点的匹配, 以找到图像间准确的匹配点对。最后, 对不同光照条件、焦距、拍摄角度获取的图像进行特征点的提取及匹配。实验结果显示, 本文算法对图像的光照、平移、旋转变换具有很好的适应性和准确性, 能够提升匹配的自动化水平和准确度, 准确率超过90%。结果表明, 提出的算法可以进一步应用到图像识别, 图像重建等领域。

**关键词:** 图像匹配; 特征提取; 尺度不变特征变换; SIFT特征; 置信度

**中图分类号:** TP391.4 **文献标识码:** A

## Image feature extraction and matching technology

YANG Xiao-min, WU Wei, QING Lin-bo, HUA Hua, HE Xiao-hai

(College of Electronics & Information Engineering, Sichuan University, Chengdu 610064, China)

**Abstract:** With the aim to improve the stability and reliability of image matching, the Scale Invariant Feature Transform algorithm(SIFT) is applied to image feature extraction and image matching. The SIFT can find out those feature vectors in different scale spaces and can extract image features and image description with the invariantce for scale changes and rotations and the flexibility for illumination variation and affine transformation. In this paper, the SIFT method is used to get the special point of an image and its features. Then, the features are matched with the criterion of the nearest neighbor based on confidence. Finally, the features in the image with different illumination conditions, focus lengths and shooting angles are extrated and matched. The experimental results prove that the features extracted by SIFT method have excellent adaptive and accurate characteristics for the illumination, transfer and the rotation transform and have match accuracy more than 90%, which are useful for the fields of image recognition and image reconstruction.

**Key words:** image matching feature extraction; Scale Invariant Feature Transform(SIFT); SIFT feature; confidence

## 1 引言

图像匹配技术是实现图像融合、图像校正、图像镶嵌以及目标识别与跟踪的关键步骤之一,已经广泛应用于图像识别以及图像重建等领域中。如何自动提取稳定可靠的特征,提高匹配准确度是实现图像匹配的重要环节,国内外学者已在这方面做了大量的研究工作。图像匹配技术主要分为基于像素的方法和基于特征<sup>[1-6]</sup>的方法。基于像素的方法直接计算两幅图像之间像素灰度的差值,认为相同部分的像素灰度差值应该比不同部分的差值要小。这种方法简单直观,但是计算量很大,而且在图像有较大旋转,较大光照、色彩差异等情况下往往会失败。为了克服基于灰度图像匹配方法的缺点,人们提出了基于特征的方法,基于特征的方法首先从待匹配的图像中提取含有图像重要特征的特征点,用相似性度量进行图像间特征点的匹配。图像匹配中常用的特征有边缘、轮廓、直线、角点等。基于特征的匹配对于图像畸变、遮挡等具有一定的鲁棒性,但是它的匹配性能在很大程度上取决于特征提取的质量,并且这一类方法的特征提取易受噪声的影响。总的来说由于成像场景等复杂条件以及不同传感器成像机理不同的特点,使得提取图像稳定的特征点变得非常困难。

尺度不变特征变换(Scale Invariant Feature Transform, SIFT)是一种基于特征的特征检测描述方法。SIFT<sup>[6]</sup>算法具有放缩、旋转和仿射不变性,能够抗拒一定光照变化和视点变换。同时,这种特征还具有较高的辨别能力,有利于后续的匹配。

本文将 SIFT 算法应用于图像特征点的提取,提取具有放缩不变性、旋转不变性,还有一定的抗

光照变化和抗视点变换的特征点,然后采用基于置信度的匹配算法进行特征点的匹配。对不同光照条件、焦距、拍摄角度的图像进行了实验,结果表明采用本文算法能够大大提升匹配的自动化水平和准确度,为图像匹配技术提供了一种新的思路。

## 2 SIFT 算法

SIFT 的主要思路是:首先建立图像的尺度空间表示,然后在尺度空间中搜索图像的极值点,由极值点建立特征描述向量。采用 SIFT 方法提取的图像特征具有放缩不变性、旋转不变性,还有一定的抗光照变化和抗视点变换性能。

### 2.1 尺度空间极值提取

尺度空间极值提取的任务就是将图像用多尺度空间表示,然后查找每一个尺度空间中的极值点,得到极值点所在像素的位置。

Lindeberg<sup>[6]</sup>等人证明高斯卷积核是唯一的线性核。在图像中常用的二维高斯函数定义如下:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right), \quad (1)$$

其中  $G(x, y, \sigma)$  是均值为 0, 方差为  $\sigma^2$  的正态分布。数字图像处理中,不同尺度下的尺度空间可以由图像与高斯函数进行卷积得到:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \quad (2)$$

其中  $I(x, y)$  为待处理的图像,  $\sigma$  是高斯函数的标准差,这里也可以理解为尺度空间的空间尺度因子。

将待处理的图像全部转换为灰度图像,使用式(1)和式(2)对图像做高斯卷积处理,在获得卷积后图像  $L$  后,使用高斯差函数(Difference-of-Gaussian, DOG)建立尺度空间:

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma), \quad (3)$$

$k$  为固定的系数。式(3)表明计算高斯差图像  $D$  时,可以先对图像进行不同尺度的高斯卷积,然后相减即可。

高斯卷积图像两两相减得到高斯差图像后,即可进行求极值的过程。将高斯差图像中的每个像素与它的邻域、它对应上一层图像的邻域、对应下一层图像的邻域一共 26 个像素点比较,如其灰度值为最大或者最小,则记录此点的位置和它所在的尺度,作为候选特征点,如图 1 所示。

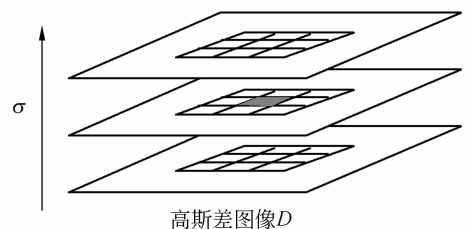


图 1 极值点查找

Fig. 1 Search of extreme point

图 2 是对一个办公桌照片做高斯差图像的例子,采用  $s=3$ ,共 5 个尺度。

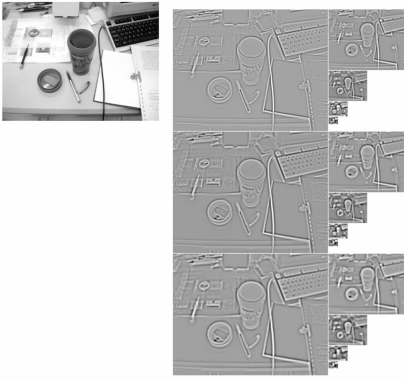


图 2 高斯差图像  
Fig. 2 DOG image

## 2.2 亚像素精度

在尺度空间极值提取步骤中已经取得了以整数形式表示的候选特征点坐标位置,需要进一步

$$D(\mathbf{x}_{\max}) = D(\mathbf{x}) + \frac{\partial D^T}{\partial \mathbf{x}}(\mathbf{x}_{\max} - \mathbf{x}) + \frac{1}{2}(\mathbf{x}_{\max} - \mathbf{x})^T \frac{\partial^2 D}{\partial \mathbf{x}^2}(\mathbf{x}_{\max} - \mathbf{x}) = D(\mathbf{x}) + \frac{\partial D^T}{\partial \mathbf{x}} \hat{\mathbf{x}} + \frac{1}{2} \hat{\mathbf{x}}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \hat{\mathbf{x}}, \quad (4)$$

其中,  $\mathbf{x} = [\sigma, x, y]^T$ ,  $D$  为高斯差图像,对式(4)求导并令其等于 0 即可解出相对极值的偏移  $\hat{\mathbf{x}}$ :

$$\hat{\mathbf{x}} = \frac{\partial^2 D^{-1} \partial D}{\partial \mathbf{x}^2 \partial \mathbf{x}}. \quad (5)$$

从而得到极值点  $\mathbf{x}_{\max}$  的亚像素位置。当偏移值  $\hat{\mathbf{x}}$  大于 0.5 时( $x$  方向或者  $y$  方向)说明更靠近另一侧的像素点,此时让另一侧像素成为候选特征点,然后重复上述计算,得到新的亚像素精度位置。最后,使用亚像素精度位置替换所有尺度的候选特征点的位置。

## 2.3 候选特征点筛选

极值点  $\mathbf{x}_{\max}$  的灰度值可以用来判断特征点是否处于低对比度的地方,低对比度容易受到噪声的干扰,这样的情况下特征点不稳定,应该舍去。将偏移值  $\hat{\mathbf{x}}$  代入式(4)可以得到极值点  $\mathbf{x}_{\max}$  的灰度值:

$$D(\mathbf{x}_{\max}) = D(\mathbf{x}) + \frac{1}{2} \frac{\partial D^T}{\partial \mathbf{x}} \hat{\mathbf{x}}. \quad (6)$$

本文将  $|D(\mathbf{x}_{\max})| < 0.008$  时的候选特征点舍去。

高斯差函数还会在边缘处产生很强的响应,需要去掉那些在边缘上容易受噪声影响的候选特征点。边缘上容易受噪声影响的位置一般具有较

确定特征点的亚像素精度,即获得以浮点数表示的精确位置。实验表明获得亚像素精度后的特征点在匹配性能和稳定性上都有显著的提高。

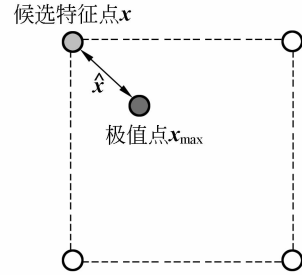


图 3 亚像素精度极值点位置  
Fig. 3 Extreme point with sub-pixel accuracy

如图 3 所示,设极值点为  $\mathbf{x}_{\max}$ ,它与候选特征点的距离为  $\hat{\mathbf{x}}$ ,  $D(\mathbf{x}_{\max})$  在邻域中一定为极值,把  $D(\mathbf{x}_{\max})$  在  $\mathbf{x}$  处泰勒展开:

大的主曲率,但在垂直方向上的曲率又较小。主曲率可以使用的 Hessian 矩阵计算:

$$H = \begin{bmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial x \partial y} \\ \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial y^2} \end{bmatrix}. \quad (7)$$

$H$  的特征值与  $D$  的主曲率成正比。设  $\alpha$  为较大的那个特征值,  $\beta$  为较小的特征值,设  $\gamma$  为  $\alpha$  和  $\beta$  的比值,即  $\alpha = \gamma\beta$ ,可以仅用  $\gamma$  来表示它们的比值:

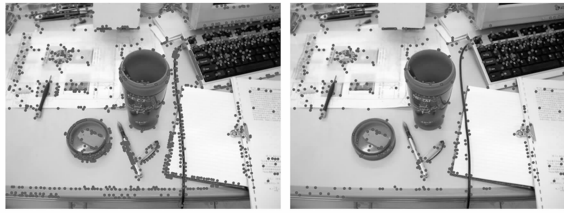
$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\gamma\beta + \beta)^2}{\gamma\beta^2} = \frac{(\gamma + 1)^2}{\gamma}. \quad (8)$$

当两个特征值相等时,式(8)获得最小值,两个特征值差别越大时其值越大。因此,对于一个给定的阈值  $\tilde{\gamma}$ ,仅仅需要检查:

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(\tilde{\gamma} + 1)^2}{\tilde{\gamma}}. \quad (9)$$

本文阈值  $\tilde{\gamma}$  取为 20。

图 4 表示了筛选前后的候选特征点位置和个数的关系,图 4(a)中共有 1 033 个候选特征点,图 4(b)共有候选特征点 461 个,可见有 50% 以上的候选特征点被删除。



(a) 所有候选特征点 (b) 筛选后的候选特征点  
(a) All feature points (b) Selected feature points

图 4 提取出的候选特征点

Fig. 4 Extracted feature points

$$\begin{cases} m(x,y)=\sqrt{[L(x+1,y)-L(x-1,y)]^2+[L(x,y+1)-L(x,y-1)]^2} \\ \theta(x,y)=\arctan \frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \end{cases} \quad (10)$$

为了凸显候选特征点的作用,在邻域上使用高斯窗函数进行加权。获得对应的梯度幅度大小和梯度方向。梯度方向的度数范围为 $[0, 360^\circ]$ ,把此范围划分为 36 个区间,每个区间  $10^\circ$ ,则形成了一个直方图。直方图中值最大的区间代表了候选特征点局部邻域中梯度的主要方向,最大值代表了这个方向上梯度的幅度大小。

### 2.5 提取特征点描述

首先,使用式(10)计算特征点对应的高斯卷积图像  $L$  中周围像素的梯度幅度大小和梯度方向。同样使用高斯窗函数对邻域梯度幅度大小进行加权。此处采用  $w/2$  作为高斯窗函数的标准差,其中  $w$  为特征点的邻域半径,在图 5 中用圆圈表示。把特征点邻域范围划分为  $4 \times 4$  的像素块,如图 5 左边所示。在每个  $4 \times 4$  的像素块中建立梯度方向直方图,此处划分为 8 个区间,并统计  $4 \times 4$  像素块中每个像素梯度方向属于的区间,这个区间的值为所有梯度方向属于这个区间的像素的梯度幅度大小高斯加权后的累加,在右边使用箭头长度表示。因此每个  $4 \times 4$  像素块都可以用一个长度为 8 维的特征描述向量表示,每一维对应于直方图的一个区间。本文在特征点周围统计  $16 \times 16$  邻域范围的像素用以描述其特征点的特征,生成  $16(4 \times 4)$  个直方图,每个直方图包含 8 个特征。所以对于每一个特征点,一共使用  $4 \times 4 \times 8 = 128$  维的特征描述向量对其进行描述,如图 5 右边所示。

记特征描述向量为  $\mathbf{N} = [n_1, n_2, \dots, n_{128}]^T$ ,

### 2.4 候选特征方向赋值

特征方向赋值就是对每一个候选特征点,通过计算它们所在局部位置的信息给它们赋予一个主方向。

首先获得候选特征点对应的高斯卷积图像,这样可以保持放缩不变性。对一个候选特征点,可以通过它的四邻域得到它所在位置的梯度幅度大小  $m(x,y)$  和梯度方向  $\theta(x,y)$ :

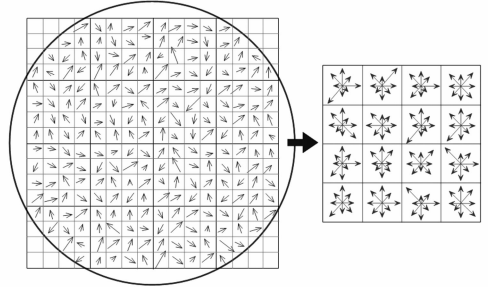


图 5 提取特征点描述

Fig. 5 Description of feature points

为了使其具有光照不变性,将特征描述向量归一化为单位长度,方法如下式:

$$n_i' = \frac{n_i}{\left(\sum_{i=1}^{128} n_i^2\right)^{\frac{1}{2}}}, i = 1, 2, \dots, 128. \quad (11)$$

### 3 匹配过程

特征点匹配是指在找出图像的特征点后,寻找图像间特征点的对应关系。通常采用最近邻方法,即查找每一个特征点在另外一幅图像中的最近邻。理想状态下两幅图像之间相同部分的特征点应该具有相同的特征描述向量,所以它们之间的距离应该最近<sup>[7-9]</sup>。但是由于图像之间存在不重叠的部分,这些区域中的特征点在另一幅图像中并没有与之对应的特征点存在,但是最近邻算法仍然会给出一个相对的最近邻做为它的匹配

点,这些匹配点对明显是错误的。为了排除这些错误的匹配点对,本文提出采用基于置信度的算法对其进行匹配,最后实验中取得了很好的效果。

置信(Confidence)广义上讲指的是正确的概率<sup>[10-11]</sup>,置信度(Confidence measures)是评价这个概率的一种量度,表示某一事件的可靠程度。由于置信度模型可以用来判断观测数据和标准模型之间的匹配程度,因此可以用来对匹配结果进行验证,从而提高整个系统的识别率和稳健性。

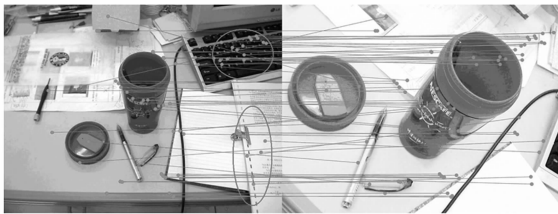
设特征点的描述向量为  $\mathbf{x}$ ,对其匹配判断的置信度为  $\text{conf}(\mathbf{x})$ ,本文定义置信度公式为

$$\text{conf}(\mathbf{x}) = 1 - d_1(\mathbf{x})/d_2(\mathbf{x}), \quad (12)$$

其中,  $d_1(\mathbf{x})$  是特征点与另一幅图像中特征点的最近距离,  $d_2(\mathbf{x})$  是特征点与另一幅图像的中特征点的次近距离。

如果  $\text{conf}(\mathbf{x})$  小于一定阈值则极有可能是一个错误的匹配点对,则将错误的匹配点对消除。

图 6 为在同一场景下不同光照、拍摄角度以及焦距条件下的图像间匹配的结果图。图 6(a) 是传统的最近邻匹配结果图的结果,图中圆圈部分为另外一幅图中没有出现的地方,所有在此处的匹配点对都是错误的。图 6(b) 为使用本文算法最后的匹配结果图,可以看出使用本文算法去除了错误的匹配点对,得到的匹配关系全部正确。



(a) 最近邻匹配结果图

(a) Nearest neighbor algorithm results



(b) 本文算法结果图

(b) Proposed algorithm results

图 6 匹配结果图

Fig. 6 Match results

## 4 实验分析

分别针对同一场景不同光照条件、拍摄角度、焦距的的光学图象,采用本文算法进行自动匹配,测试匹配效果。

### 4.1 检测算法对光照变化的适应能力

分别采集了同一场景不同光照条件下的图像,检验本文算法对于光照变化的适应能力。结果如图 7 所示。



图 7 不同光照条件下图像间的匹配结果图

Fig. 7 Match results between the two images at different illumination conditions

### 4.2 检测算法对旋转变化的适应能力

分别采集了同一场景不同角度下的图像,检验本文算法对旋转变化的适应能力。结果如图 8 所示。



图 8 不同拍摄角度下图像间的匹配结果图

Fig. 8 Match results between the two images at different shooting angles

### 4.3 检测算法对放缩变化的适应能力

分别采集了同一场景不同焦距条件下的图



图 9 不同焦距条件下图像间的匹配结果图

Fig. 9 Match results between the two images at different focal lengths

像,检验本文算法对于缩放的适应能力。结果如图9所示。

#### 4.4 检测算法对旋转、放缩和光照都发生改变的适应能力

分别采集了同一场景不同光照条件、不同拍摄角度以及不同焦距条件下的图像,进行自动匹配,来测试算法效果,结果如图10所示。



图10 不同光照、拍摄角度、焦距下图像间的匹配结果图

Fig. 10 Match results between the two images at different focal lengths, shooting angles and illumination conditions

#### 4.5 本文算法与传统匹配算法比较

分别针对不同场景在不同光照条件、不同角度以及不同焦距条件下采集图像(实验样本如图11所示)进行特征点提取与匹配,并将本文算法与传统算法进行比较,结果如表1所示。

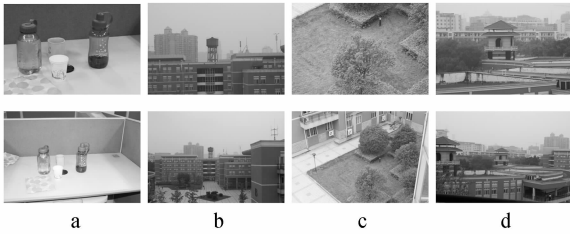


图11 实验样本图  
Fig. 11 Image samples

#### 参考文献:

- [1] KELLER Y, AVERBUCH A, ISRAELI M. Pseudo-polar based estimation of large translations rotations and scalings in images[J]. *IEEE Transaction on Image Processing*, 2005, 14(1): 12-22.
- [2] SIGGELKOW S. *Feature histograms for content-based image retrieval*[D]. Frieiburg: Albert Ludwigs University of Frieiburg, 2002.
- [3] MIKOLAJCZYK K, SCHMID C. An affine invari-

表1 算法比较结果

Tab. 1 Comparison of different algorithms

不同组图片	Harris 角点法	基于像素的算	本文算法
	匹配正确率	法匹配正确率	匹配正确率
a 组图片	60%	57%	98%
b 组图片	79%	61%	91%
c 组图片	76%	47%	92%
d 组图片	83%	52%	96%

由实验可以看出, Harris 角点方法对噪声比较敏感,随着图像分辨率的不同,角点容易产生漂移,因此对于不同尺度的同一场景的图片(a组图片)效果并不理想。基于像素的方法在图像光照、色彩差异等情况(b,c,d组图片)下往往效果都很差。本文算法对图像的光照、平移、旋转变换的情况都具有很好的适应性和准确性。

## 5 结论

在图像特征提取与匹配领域中,提取稳定可靠的特征是提高匹配准确度的一个重要环节。本文采用 SIFT 算法提取图像的特征点,其效果不受图像伸缩、旋转、光照的影响,能够正确地提取出描述物体本质的信息;然后采用基于置信度的匹配算法进行特征点的匹配,该算法保留了正确的 SIFT 特征匹配点,并有效地剔除误配点,显著地改善了匹配结果,实现图像间特征点之间的准确匹配。通过对不同光照条件、焦距、拍摄角度的图像进行一系列的实验,实验结果表明本文算法对图像的光照、平移、旋转变换具有很好的适应性和准确性,并能够大大提升匹配的自动化水平和准确度,准确率可达 90% 以上,可以进一步应用到图像识别,图像重建等领域。

ant interest point detector[C]. *Proceedings of the 7th European Conference on Computer Vision*, 2002, 128-142.

- [4] ZHANG Z Y, DERICHE R, FAUGERAS O. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry [J]. *Artificial Intelligence*, 1995, 78 (2): 87-119.
- [5] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of*

*Computer Vision*, 2004, 60(2):91-110.

- [6] LINDBERGER T. Feature detection with automatic scale selection [J]. *International Journal of Computer Vision*, 1998, 30(2):79-116.
- [7] 王会峰, 刘上乾, 汪大宝, 等. 基于序列图像特征匹配的摄像机旋转补偿算法[J]. *光学精密工程*, 2008, 16(7):1330-1334.  
WANG H F, LIU SH Q, WANG D B, *et al.*. Video camera rotation compensation algorithm based on feature matching of sequence image frames[J]. *Opt. Precision Eng.*, 2008, 16(7):1330-1334. (in Chinese)
- [8] 罗诗途, 王艳玲, 张玘, 等. 车载图像跟踪系统中电子稳像算法的研究[J]. *光学精密工程*, 2005, 13(1):95-103.  
LUO SH T, WANG Y L, ZHANG Q, *et al.*. Electronic image stabilizing algorithm for image tracking

system on vehicle[J]. *Opt. Precision Eng.*, 2005, 13(1):95-103. (in Chinese)

- [9] 丁雪梅, 王维雅, 黄向东. 基于差分 and 特征不变量的运动目标检测与跟踪[J]. *光学精密工程*, 2007, 15(4):570-576.  
DING X M, WANG W Y, HUANG X D. New method for detecting and tracking of moving target based on difference and invariant[J]. *Opt. Precision Eng.*, 2007, 15(4):570-576. (in Chinese)
- [10] KITTLER J, HATEF M, DOIN P R W, *et al.*. On combining classifier [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(3):226-239.
- [11] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient belief propagation for early vision [J]. *International Journal of Computer Vision*, 2006, 70(1):41-54.

#### 作者简介:



杨晓敏(1980—),女,四川广安人,讲师,博士,2002年于四川大学获得学士学位,2007年于四川大学获得博士学位,主要从事图象处理和图象通信方面研究。E-mail: arielyang2006@163.com



华 骅(1981—),男,四川成都人,工程师,硕士,2003年于四川大学获得学士学位,2007年于四川大学获得硕士学位,主要从事图像处理和模式识别研究。E-mail: doubhua@gmail.com



吴 炜(1975—),男,四川宜宾人,讲师,博士,2003年于四川大学获得硕士学位,2008年于四川大学获得博士学位,主要从事图像处理和模式识别研究。E-mail: wuwei75@yeah.net



何小海(1964—),男,四川盐亭人,教授,博士生导师,2003年于四川大学获得硕士学位,2008年于四川大学获得博士学位,主要从事图像处理和图像通信方面研究。E-mail: chenmo1984@gmail.com



卿 澜波(1982—),男,四川简阳人,讲师,博士,2003年于四川大学获得学士学位,2008年于四川大学获得博士学位,主要从事图像处理和模式识别研究。E-mail: qing\_lb@scu.edu.cn