

基于极坐标图像投影直方图与FFT的象棋识别方法

胡鹏¹, 罗杨宇¹, 李成荣¹

1. 中国科学院自动化研究所, 北京 100190

E-mail: phu@hitic.ia.ac.cn

摘要: 基于机器视觉的中国象棋的识别包括棋子的检测和棋子字符的识别。为达到快速、鲁棒的系统性能要求, 采用Robert算子提取棋子的边缘轮廓, 然后通过形态学处理、模板圆方法对棋子进行检测、定位和分割。在特征提取方面, 提出了极坐标图像投影直方图与快速傅立叶变换相结合的方法, 快速提取字符的旋转不变性特征。实践证明, 该方法处理速度快、识别率高、抗噪能力强。

关键词: 象棋识别, 旋转字符识别, 极坐标图像, 投影直方图, 快速傅立叶变换

Chinese Chess Recognition based on Projection Histogram of Polar Coordinates Image and FFT

Peng Hu¹, Yangyu Luo¹, Chengrong Li¹

1. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

E-mail: phu@hitic.ia.ac.cn

Abstract: The recognition of Chinese chess through computer vision includes chess detection and its character recognition, and the process should be fast and robust. In this paper, Robert operator is used to get the edge information of chess, then detection, localization and segmentation of the chess are realized through mathematical morphology and template circle method. A new algorithm based on projection histogram of polar coordinates image and Fast Fourier Transform is introduced to extract the rotation-invariant feature of chess characters. Experiments show that the algorithm can accurately recognize all chesses within 300 ms, and is robust to any rotation.

Key Words: Chinese chess recognition, rotation characters recognition, polar coordinates image, projection histogram, Fast Fourier Transform

1 引言

象棋是中国流传千年之久的智力游戏, 是中华文化的精粹之一。象棋机器人的研制在国内开始慢慢多起来, 它是机械设计、人工智能、自动控制、计算机视觉等方面理论与技术的综合应用。本文介绍象棋机器人的视觉部分, 从CCD摄像机拍摄的棋盘图片中检测、识别棋子, 然后将识别结果提供给下棋软件进行推理。

象棋识别的关键步骤为棋子的检测和棋子字符的识别。针对棋子的检测, 文献[1]、[2]采用相邻像素灰度差阈值的图像二值化方法对图像进行预处理, 然后用模板圆按棋盘的先验位置检测棋子。文献[3]、[4]将棋盘图像按棋子的先验位置进行分块, 然后采用自适应阈值方法二值化子图像, 再用Hough变换检测圆, 从而提取出棋子。利用棋子先验位置信息进行棋子检测的方法只能处理棋盘摆放无旋转、无倾斜以及棋子摆放位置准确的情况, 采用各种阈值方法对图像进行二值化处理就要求棋子与棋盘的颜色差异很大, 而且光照均匀, 无阴影。

这些方法只能处理特定约束条件下的棋盘图像, 不具有普适性。在棋子字符识别方面, 文献[2]提出了统计文字连通数与孔数、年轮统计和基于图像配准的模板匹配等方法。文献[3]统计字符的同心圆环内的像素点数, 并采用神经网络的方法进行训练识别, 文献[4]在文献[3]的基础上采用SVM作为分类器进行识别。这些特征都不够稳定, 当棋子字符因部分反光或者对焦模糊等情况发生时, 二值化后的字符笔画中可能有孔洞、边缘残缺、笔画粘连等问题, 这些特征就都失效了。

本文采用Robert算子提取象棋图像边缘, 然后形态学方法加以处理, 用模板圆快速检测出棋子。特征提取方面, 将字符图像变换为极坐标图像, 然后统计竖直投影直方图, 采用快速傅立叶变换(FFT)去除直方图的平移性。最后通过计算模板特征向量与样本特征向量的相似度进行字符识别。本文算法简洁, 速度快, 识别率高, 鲁棒性强。本文的结构介绍, 第一部分为引言, 第二部分为棋子的定位, 第三部分为棋子识别, 最后为结论部分。

2 棋子检测

一般象棋棋子上的字符为红色和绿色或者红色和黑色。利用颜色分割图像是一种很直接的方法，能快速的分离出棋子。但颜色信息受环境光照影响很大，而且每次都需提取先验颜色。要想更有效、更鲁棒地检测出棋子，还是应该利用纹理信息。字符的边缘信息丰富，而且棋子与棋盘的灰度差异很大，边缘轮廓明显，所以利用边缘图检测棋子不失为一个很好的方法。

2.1 边缘提取

边缘提取的方法有很多，也很成熟，包括Canny算子、Sobel算子、Laplace算子、Robert算子等等。经过实验对比发现，简单的Robert算子提取的棋盘边缘图最利于棋子的检测。

Robert算子是最早的边缘提取算法之一，通过计算相邻像素的对角差分获得图像的边缘图。计算公式如2.1.1，公式中将阈值分割也包含在内，当差分值大于阈值 $thresh$ 时，对应的像素置1。

$$f(x,y)=\begin{cases} 1 & \text{if } (|p(x,y)-p(x+1,y+1)|+|p(x+1,y)-p(x,y+1)|>thresh) \\ 0 & \text{else} \end{cases} \quad 2.1.1$$

原图及效果图如图2.1.1和图2.1.2所示。

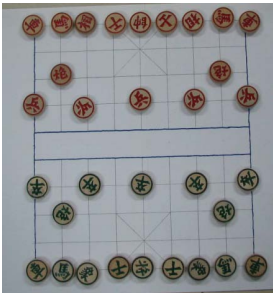


图2. 1. 1 原图

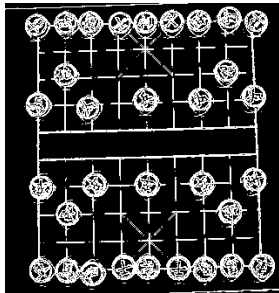


图2. 1. 2 Robert边缘图

为便于检测圆域，需用形态学方法对边缘图进行处理。为填充区域的内部空洞，采用膨胀算子运算，效果如图2.1.3所示。

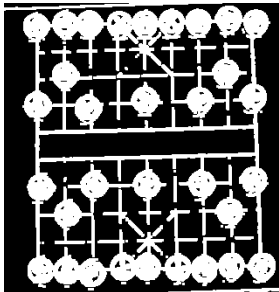


图2. 1. 3 膨胀运算结果

2.2 模板圆检测棋子

对于圆域轮廓的检测，最常用的方法为模板圆法和Hough变换。模板圆需提供半径信息，而Hough变换没有此要求，当然Hough变换也可以在确定半径的情况下对圆心位置进行投票。但Hough变换计算量大，而且对于象棋来说，棋子有阴影干扰，并且有可能粘连在一块，使得Hough变换检测的结果很差。本文采用模板圆的检测方法，对于每个目标点，统计其模板圆域的像素点数，然后检测出拥有局部最大值的点。

模板圆域的像素点数统计公式如式2.2.1， f 为膨胀运算后的结果， r 为模板圆半径。

$$s(x,y)=\sum_{i,j} f(x+i,y+j), \quad (i^2+j^2 \leq r^2) \quad 2.2.1$$

采用阈值滤掉 s 值较小的点，得到部分候选圆心点如图2.2.1所示，然后求取局部极大值点，即正确的棋子圆域，如图2.2.2所示。

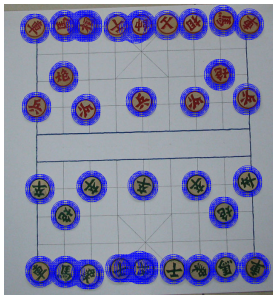


图2. 2. 1 部分候选圆域



图2. 2. 2 求取局部极大后的结果

3 棋子识别

棋子识别是一个典型的旋转字符识别问题。识别的关键是要提取字符的旋转不变性特征向量。传统的方法中理论相对成熟的有Hu矩特征[6]、Zernike矩特征[7]，还有先提取字符外轮廓点，然后采用傅立叶变换的方法[8][9]等等。在针对棋子的识别中，也有研究者提出年轮法、连通数与孔数法等。经大量实验发现这些方法都不够鲁棒，抗干扰能力差，且不同字符之间的特征差异不大。因此本文提出一种新的针对旋转字符的特征提取方法，该方法将图像转换为极坐标图像，然后采用经典的投影直方图方法，再利用FFT去除直方图的平移性。

3.1 棋子预处理

分割出来的棋子需要进行预处理才方便提取特征。采用OTSU方法二值化后的棋子有粗糙的圆框（图3.1.1），会给识别带来干扰信息，本文利用像素点分布的特征对圆域上的像素进行若干采样，然后用RANSAC方法估计出圆的参数，成功的去掉边框（图3.1.2）。



图3.1.1 二值图



图3.1.2 去边框后

3.2 坐标变换

图像在笛卡儿坐标系下的旋转变换对应于极坐标系下的平移变换，关系如式3.2.1所示， f_r 为原图像 f 旋转 θ_0 后的新图像。在3.1中估计出圆框的圆心，以圆心为坐标原点，将笛卡儿坐标图像转换为极坐标图像，效果如图3.2.1所示。

$$\begin{cases} f_r(x, y) = f(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0) \\ f_r(\rho, \theta) = f(\rho, \theta - \theta_0) \end{cases} \quad 3.2.1$$

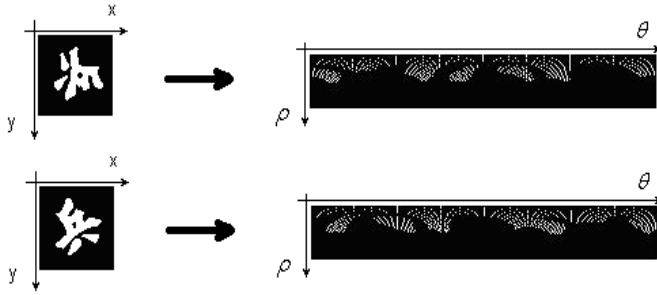


图3.2.1 坐标变换

对极坐标图像的横竖两个方向统计投影直方图等价于在笛卡儿坐标图像的径向和切向（扇面与同心圆环）统计投影直方图。实验表明坐标变换是必要的，直接在笛卡儿坐标图像中统计投影直方图，会由于坐标离散化的原因而漏掉部分数据未进行统计。坐标变换中，原图像中的每一个点都映射到新的极坐标图像中，信息全部保留了下来参与特征的计算。

3.3 投影直方图

针对笛卡儿坐标图像的横竖两个方向的投影直方图统计很早就被作为一种成熟的技术用在文本检测、识别中，常用来分割文本段落、文本行、字符等，有时候也用于小数据集的字符识别。虽然它是一种容错性好的统计特征，但对字符倾斜比较敏感，而且丢失了大部分的空间结构信息，使得它对字符的特征表达能力不强。然而，在极坐标图像中，字符的空间结构进行了切向展开，极坐标图像的投影直方图就保留了字符的大部分结构信息。

投影直方图的表述公式为：

$$\begin{cases} n_\theta(\theta) = \sum_\rho f_r(\rho, \theta) \\ n_\rho(\rho) = \sum_\theta f_r(\rho, \theta) \end{cases} \quad 3.3.1$$

$n_\rho(\rho)$ 对字符的表征能力相当弱，所以本文只取 $n_\theta(\theta)$ 进行字符识别（图3.2.2）。如果极坐标图像的像素点很稀疏，可以考虑先将其进行一次形态学膨胀操作，让像素点变得更密集，然后再统计直方图。

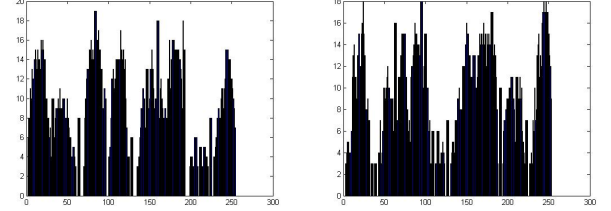


图3.2.2 两个不同角度的“兵”的极坐标直方图

3.4 快速傅里叶变换 (FFT)

傅立叶变换是对信号进行频域分析的方法之一，快速傅立叶变换 (FFT) 是一种高效的离散傅立叶变换 (DFT) 计算方法。

离散傅立叶变换的表述为：

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi j}{N} nk} \quad 3.4.1$$

快速傅立叶变换有很多种形式，使用最广泛的是 Cooley-Tukey FFT 算法[10]，它是将 DFT 分解成奇偶两部分进行 DFT，递归的分解下去，直到每个小 DFT 只有两项的和式计算。表述形式如下：

$$\begin{aligned} X_k &= \sum_{m=0}^{\frac{N}{2}-1} x_{2m} e^{-\frac{2\pi j}{N} (2m)k} + \sum_{m=0}^{\frac{N}{2}-1} x_{2m+1} e^{-\frac{2\pi j}{N} (2m+1)k} \\ &= \sum_{m=0}^{\frac{N}{2}-1} x_{2m} e^{-\frac{2\pi j}{M} mk} + e^{-\frac{2\pi j}{N} k} \sum_{m=0}^{\frac{N}{2}-1} x_{2m+1} e^{-\frac{2\pi j}{M} mk} \\ &= \begin{cases} E_k + e^{-\frac{2\pi j}{N} k} O_k & \text{if } (k < M) \\ E_{k-M} - e^{-\frac{2\pi j}{N} (k-M)} O_{k-M} & \text{if } (k \geq M) \end{cases} \end{aligned} \quad 3.4.2$$

信号在时域的平移性对应于频域的旋转性，只有相频变化，而不影响幅频。表述形式为：

$$\begin{aligned} f(x) &\rightarrow F(\omega) \\ f(x+t_0) &\rightarrow F(\omega) e^{-j\omega t_0} \end{aligned} \quad 3.4.3$$

因此，针对极坐标图像的竖直投影 $n_{\theta}(\theta)$ ，要去掉其平移性，就需对其进行傅立叶变换，然后取幅频信息，得到最终的旋转不变性特征向量 V 。

$$V = |F[n_{\theta}(\theta)]| \tag{3.4.4}$$

3.5 识别

模式识别问题都可分为两个部分：特征的提取和特征的识别。表征能力很强的特征与弱的分类器组合有可能达到很好的识别效果，粗糙的特征与强分类器组合也可能达到很好的识别效果。

由于本文基于极坐标图像投影直方图和快速傅立叶变换提取的特征对字符的表达能力强，所以用简单的相似度函数就可以达到100%的识别正确率。度量两个向量的相似性最常用的函数是两向量的夹角余弦，如式3.5.1。

$$S(V_1, V_2) = \cos \angle V_1, V_2 = \frac{V_1 \cdot V_2}{|V_1| |V_2|} \tag{3.5.1}$$

当然如果采用人工神经网络或者SVM训练出一个分类器对特征进行识别，识别效果肯定更加可靠。

最后是确定棋子在整个棋盘中的相对位置。可以根据棋盘的边框四个角点计算出单应矩阵 H ，然后将棋子在棋盘中的坐标位置一一确定。实验结果如图3.5.1和3.5.2所示。

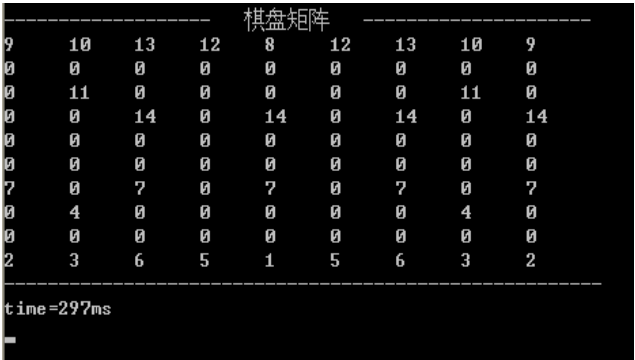


图3.5.1 控制台程序的终端输出结果

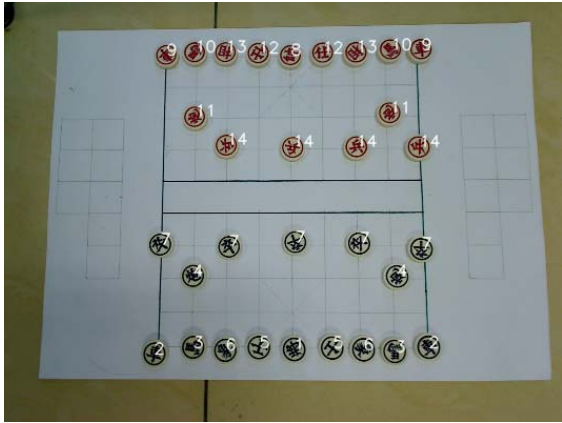


图3.5.2 棋子识别结果

实验所用摄像头为罗技快看合金版Pro9000，图像大小为1600×1200，总的计算时间为297ms，正确率100%，系统配置为Intel Pentium 4 CPU 3.00GHz，512MB内存。

4 结论

本文提出的基于Robert边缘图及模板圆的棋子检测方法，能快速准确的检测出棋子。在棋子特征提取过程中，本文提出一种新的、简单的特征提取方法，该方法是在极坐标图像中统计投影直方图，然后采用快速傅立叶变换去除直方图的平移性得到稳定的特征向量。该特征提取方法理论简单，且可以推广到一般性的旋转字符识别和旋转图像识别中。由于特征鲁棒，使用简单的相似度函数就能很好的区分特征。本文设计的象棋识别系统对棋盘摆放位置没有要求，可以在不旋转图片的前提下识别任意角度的棋盘。

参考文献

[1] 杜俊俐, 黄心汉. "象棋机器人视觉系统设计." 电子技术应用 33(09): 133-136, 2007.

[2] 杜俊俐, 张景飞等. "基于视觉的象棋棋盘识别." 计算机工程与应用 43(034): 220-222, 2007.

[3] 张志伟, 孔凡让等. "对弈机器人的视觉图像处理和识别." 计算机应用与软件 25(002): 215-217, 2008.

[4] 赵吉文, 张志伟等. "基于 SVM 的仿人对弈机器人视觉图像处理." 系统仿真学报 19(018): 4235-4238, 2007.

[5] 金元郁, 李新等. "基于图像处理的人和机械手象棋对弈系统实现." 青岛科技大学学报 28(001): 73-75, 2007.

[6] Hu, M. K. X. "Visual pattern recognition by moment invariants." Information Theory, IRE Transactions on 8(2): 179-187, 2007.

[7] Khotanzad, A. and Y. H. Hong. "Invariant image recognition by Zernike moments." IEEE Transactions on Pattern Analysis and Machine Intelligence 12(5): 489-497, 1990.

- [8] Kuhl, F. P. and C. R. Giardina. "Elliptic Fourier features of a closed contour." *Computer graphics and image processing* 18(3): 236–258, 1982.
- [9] Zahn, C. T. and R. Z. Roskies. "Fourier descriptors for plane closed curves." *IEEE Transactions on computers* 21(3): 269-281, 1972.
- [10] Cooley, James W., and John W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.* 19, 297–301, 1965.