

基于图像识别的SolidWorks三维重建技术研究

毕思远, 沈景凤, 仲梁维

上海理工大学机械工程学院, 上海

收稿日期: 2024年12月27日; 录用日期: 2025年1月20日; 发布日期: 2025年1月29日

摘要

本研究聚焦于基于图像识别的三维重建技术, 强调高层理解的扫描图纸作为重要前提, 并在机械知识的引导下实现三维模型的重建。图像识别分为基本图素识别和字符识别两个部分, 其中基本图素识别采用改进的Hough变换法, 字符识别采用基于卷积神经网络的方法, 以适应实际图样的特殊情景。实验结果表明, 该模型具有显著的识别效果, 能够有效兼容干扰图线, 识别率高达98.65%。随后, 通过工程图的语义知识, 建立基本图素之间的拓扑关系和尺寸约束图形链, 实现对扫描图纸的二维重建。最终, 在机械领域知识的支持下, 充分利用图纸识别后的信息和知识推理, 实现三维形体的重建。

关键词

图像识别, 三维重建, 卷积神经网络

Research on SolidWorks 3D Reconstruction Technology Based on Image Recognition

Siyuan Bi, Jingfeng Shen, Liangwei Zhong

School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Dec. 27th, 2024; accepted: Jan. 20th, 2025; published: Jan. 29th, 2025

Abstract

This study focuses on 3D reconstruction techniques based on image recognition, with an emphasis on the high-level understanding of scanned drawings as a crucial prerequisite. The aim is to realize the reconstruction of 3D models under the guidance of mechanical knowledge. The image recognition process is divided into two parts: basic pixel recognition and character recognition. The improved Hough transform method is adopted for basic pixel recognition, while character recognition utilizes a method based on convolutional neural networks to adapt to the specific characteristics of actual drawings. Experimental results demonstrate that the model has a significant recognition

effect, effectively handling interference lines with a high recognition rate of 98.65%. Furthermore, by leveraging semantic knowledge from engineering drawings, the topological relationship between basic graphic elements and size constraint graph chains is established to achieve two-dimensional reconstruction of scanned drawings. Finally, with support from mechanical domain knowledge, 3D shape reconstruction is realized through comprehensive information utilization and knowledge reasoning following drawing recognition.

Keywords

Image Recognition, 3D Reconstruction, Convolutional Neural Network

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在产品数字化模型[1]中,从二维到三维的智能映射是通过解析工程图中的几何信息和拓扑信息,生成相应的三维形体,这一过程被称为三维重建,主要方法有两类:自底向上的重建方法和基于模型引导的重建方法[2]。

典型的自底向上方法是 Masanori 在 1973 年提出的基于边界表示模型(Boundary Representation, B-rep)[3],该算法主要采用点-边-面的分层构造思想,步骤简单、条理清晰,但容易产生不属于三维形体的假点、假边和假面;Wesley-Markowsky 算法[4]通过列举所有可能解,达到了消除假元的效果;刘世霞、Liu S 等提出了二次曲面体的三维重建[5],首次提出使用共轭直径法,节约了大量的空间和时间成本;龚洁晖等提出的斜截切二次曲面重建算法[6],通过深度优先搜索识别曲面,并利用视图外轮廓计算交线,最终生成混合线框模型来构建三维;Alexei 和 Arkady 认为线框模型能表示任意三维实体[7],基元体通过矩阵和关联矩阵表示,利用预设模式求解三维结构,适用于清晰的矢量工程图。

基于模型引导(Constructive Solid Geometry, CSG)的重建方法由 Aldefeld 提出[8];张雪、赵德胜等基于 CSG 将基元分类组合[9],利用基元之间的拓扑关系实现三维重建,解决了特征匹配差及重影问题;Chen Zen 和 Perng Der-Baau 提出基于体切削的三维重建算法[10],通过不同策略识别和构造平面,利用深度优先搜索法逐步逼近原视图完成三维构建,但无法解决病态解问题。

现有研究大多侧重于矢量工程图[11],未充分利用领域内知识和二维图纸与三维模型之间的共享关系,输出结果是一个个图元组成的矢量文件,图元之间没有关联,都是零散的图形线段,不具有实际工程意义。故本文以轴类产品为对象,对扫描图纸进行图文分离,采用基于 Hough 变换原理的改进方法对图素进行研究,结合基于卷积神经网络的字符识别方法,实现二维重建,再利用投影、尺寸与表达等机械领域的知识,按特征进行基元重建,最终完成三维重建。

2. 知识库与整体方案

知识库的建立是智能识别与映射的基础,本章对知识库进行详细的描述,并阐述了知识引导下的数字化模型智能识别与映射的整体方案。

2.1. 知识库的建立

如图 1 和图 2,首先从工程图纸中提取特征,然后在实例库中搜索子级和父级实例,最终根据规则完

成三维模型的构建。产品模型由部件组成，部件又由零件构成，形成层次分明的结构[12]。因此，首先需要对产品模型进行特征实例分解，分解依据为产品的功能和结构。

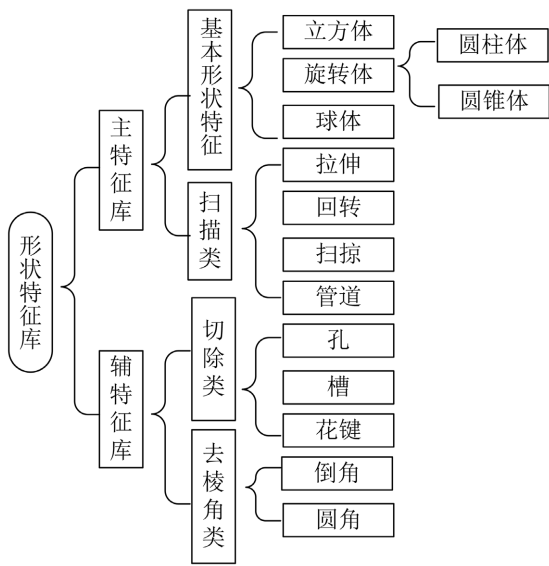


Figure 1. Classification of shape features
图 1. 形状特征的分类

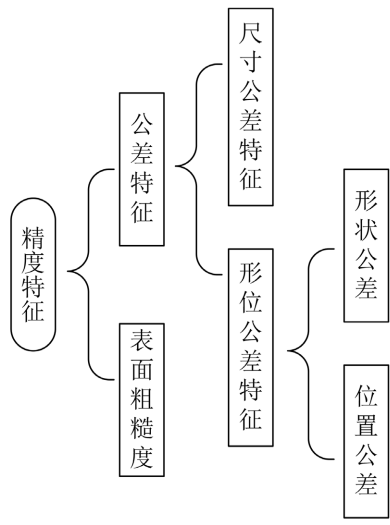


Figure 2. Classification of precision features
图 2. 精度特征的分类

2.2. 三维重建方案

图 3 为知识引导下的整体实施方案，产品数字化模型[13]通过智能识别产品结构、几何和加工信息，提取特征知识和语义信息，实现设计与重用。在知识库的基础上，首先对扫描图纸预处理及图文分离，字符和图像分别进行字符识别和基本图素识别，建立图素间的拓扑关系，并构建尺寸约束图形链，实现高层理解。完成二维几何重建后，通过投影、尺寸和机械领域知识进行基元重建，最终完成三维模型的建立。

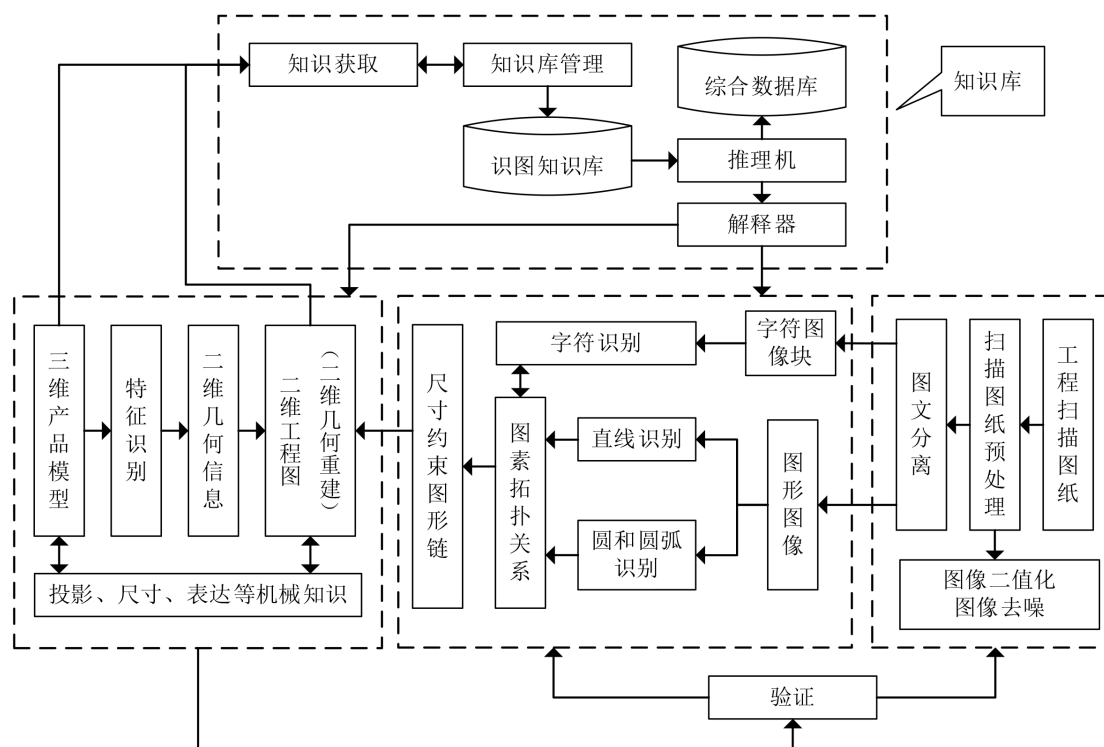


Figure 3. A holistic approach guided by knowledge

图 3. 知识引导下的整体方案

3. 基本图素识别

3.1. 图文分离

工程图由字符和图形组成, 识别时需分离两者, 以提高准确性。图形部分包含几何和拓扑信息, 字符部分则涉及语义理解。图文分离的核心是字符定位, 这是一种图像分割算法, 用于标识字符区域[14]。

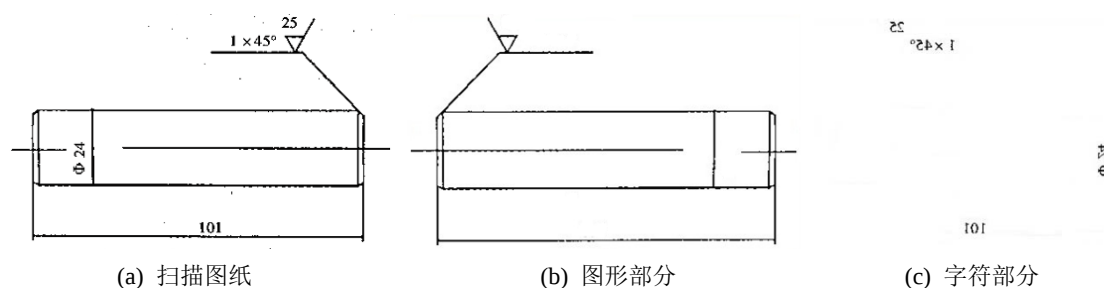


Figure 4. Flow chart of Graph separation algorithm

图 4. 图文分离示意图

一般情况下，字符的特征容易区分，但常见的特殊标注和表示方法往往不符合标准特征，且文字标注与线条粘连，难以清晰区分字符区域，导致识别精度低。为了有效分离字符与图形，首先需要对预处理后的图像进行连通域标记，识别字符区域。然后，通过设定字符特征阈值，提取字符的连通域区域，最终将图形和字符分开。原扫描图纸如图 4(a)所示，图文分离后的图形和字符信息分别如图 4(b)和图 5(c)所示。

3.2. 图像的细化

细化处理对象通常是预处理后的二值化图像，细化过程是将目标像素边界不断剥落，直至目标像素成为细线，常见的细化算法有 Hilditch 算法和查表法。

Hilditch 算法[15]是按从左至右、从上往下的顺序对图像进行遍历，如果像素满足可删除的条件，则标上标记，待像素扫描完成后，将标记的像素点用背景像素代替，再进行新的扫描。

查表法[16]是根据一定的规则来指定一个表，从而决定目标像素点是否删除，该算法实行的依据是目标像素点周围八邻域像素的色彩顺序。如图 5，(1)中从左上角开始，按顺时针方向，八邻域像素为“黑黑白白白黑”，其中 0 表示黑像素点，1 表示白像素点，则色彩顺序就变成 00111110，转换成十进制就是 62，同理(2)中对应的是 238，(3)对应的是 171，(4)对应的是 127，(5)对应的是 255，(6)对应的是 0。共有 256 种色彩顺序，分别对应 0~255 共 256 个数，这样就把色彩顺序和数字对应起来。

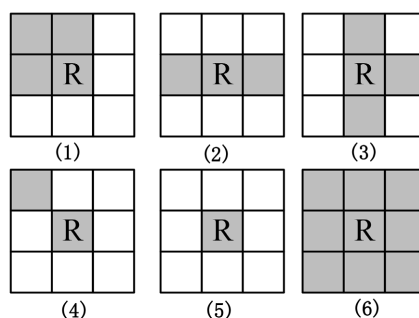


Figure 5. The octet color order of R
图 5. R 的八邻域色彩顺序

```
int[] search_array = { 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1,
    1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1,
    0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1,
    1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1,
    1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1,
    1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1,
    0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1,
    1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0,
    1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0,
    1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0,
    1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0 };
```

Figure 6. Color sequence table
图 6. 色彩序列列表

(4)中只有一个黑像素点与 R 紧邻，表明这是一条直线的两个端点，所以也不能删除；(5)表明 R 为孤立点，也不能去除；在(6)中，R 像素点周围都是黑像素点，说明 R 为内部像素点，不能去除；(1)、(2)、(3)中，R 是可以被去除的。R 像素点的色彩顺序一共有 2^8 种，由于数量太多不容易列举，故需要制作一张序列表来判断 R 像素点八邻域的 256 种色彩顺序，如图 6，该序列列表用 array 来表示。

该序列表中的元素只有 1 和 0 两种。在进行处理时首先判断像素点的状态，如果是白像素点，则不用处理；如果是黑像素点，则读取它的八邻域色彩顺序，根据八邻域的状态得到对应的索引值来判断是否删除该点，若查得的索引值是 1，说明该点可以删除，否则不能删除。由于 Hilditch 算法容易受噪声干

扰且骨架中心易发生偏移，故本文采用查表法，细化效果如图 7 所示。

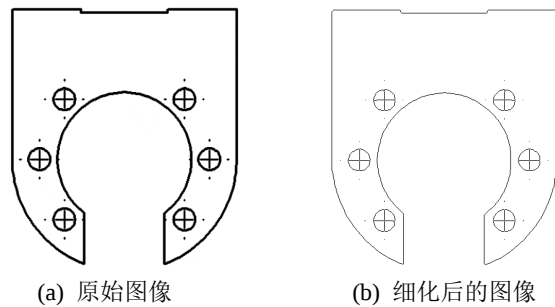


Figure 7. Schematic diagram of image refinement results
图 7. 图像细化结果示意图

3.3. 直线识别

Hough 变换是一种常用的直线检测方法[17]，通过图像空间转换到参数空间，利用直线与点的对偶关系，将直线检测问题转化为点的检测问题。该方法能有效识别边缘特征明显的图形，对噪声具有较强的鲁棒性，主要得益于通过局部度量信息获得全局参数的特点。

在 Hough 变换中，图像的每个边缘点被映射到参数空间中的一条正弦曲线。通过将极坐标系量化为小格，并对每个小格进行累加，可以得到一个累加矩阵。如图 8，在参数空间中，直线对应的局部极大值即为目标直线的参数。

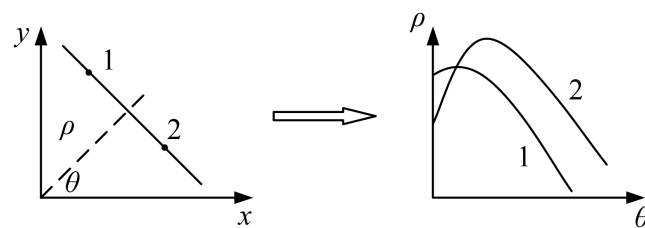


Figure 8. Hough transform in line detection
图 8. 直线检测中的 Hough 变换

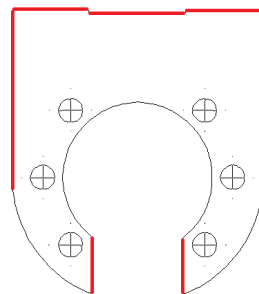


Figure 9. Schematic diagram of line recognition results
图 9. 直线识别结果示意图

该方法的优点是 ρ 和 θ 的取值不根据直线的改变而改变，但缺点是当 θ 量化的值太大时，它的运算复杂度可达到 $n(o^3)$ ，对于字节量很大的图像来说，一方面内存太大，另一方面受 θ 步长的影响，对于任意斜率的直线检测速度不够快。

若想要快速且精确地检测任意直线,就需要建立一个不受步长影响的参数空间,这里引入一种斜率分式查表法。它用分式来表示直线的斜率,分子和分母都需要是整数。有些分式简化后则表示同样的斜率,如 $4/7$ 和 $16/28$ 这两个分式,故需要对分式进行简化处理。以 800×600 的图像为例,初始化一张表格[18],该表格是将 $1 \sim 800$ 的整数组成的分式进行简化,如 $10/14$ 简化后就是 $5/7$,则数组可以记为 $\text{Array}[10].\text{sr}[14][0]=5$, $\text{Array}[10].\text{sr}[14][1]=7$, sr 为数组 Array 的数据成员,这样查找数组表格就可以快速得到化简结果。图 9 是利用斜率分式查表法识别出的直线。

3.4. 圆和圆弧识别

传统的 Hough 变换检测圆[19]与检测直线的方法类似,只是增加了未知参数个数,并且将二维的坐标平面变成三维的立体空间,如图 10, $x-o-y$ 平面内的一点 (a,b) 映射到 o_1-o_2-r 参数空间内是一个三维的锥面。

为了确定 r 、 o_1 、 o_2 这三个未知参数,需要建立一个三维的参数矩阵 $T(o_1, o_2, r)$, 形成一个带有网格计数器的三维立体网格矩阵。若 $x-o-y$ 平面内的点对应的三维曲面经过某个立体网格,那么计数器进行加 1 操作。检测完所有的立体网格后,计数器中的最大值 $T_{1_{2max}}$ 对应于 $x-o-y$ 平面内圆的半径 r 和圆心坐标 (o_1, o_2) 。传统 Hough 变换的本质是从低维到高维的映射,图像的像素越大,计算量就越大,从而效率降低。

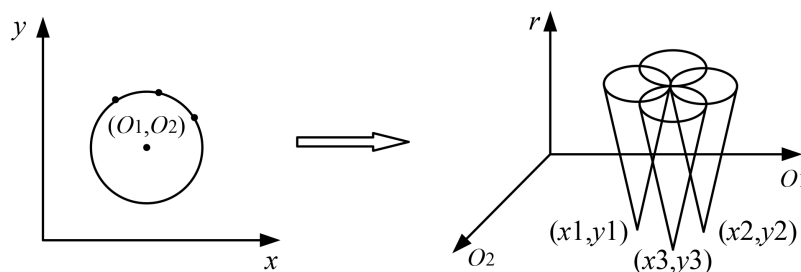


Figure 10. Hough transform detection circle diagram

图 10. Hough 变换检测圆示意图

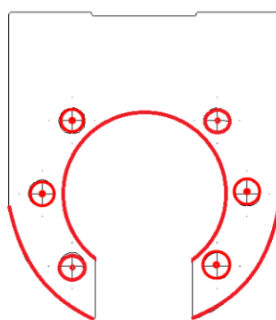


Figure 11. Circle and arc recognition results

图 11. 圆和圆弧识别结果示意图

改进的 Hough 变换检测圆方法将 Hough 变换与圆的几何特性结合起来,采用 Canny 算子对图像进行轮廓检测,使得算法复杂度降低,并且提高准确率。

将每条轮廓分为三段,并分别选择每段的三个点 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) , 根据下面的方程式求解出圆心和半径,并进行存储:

$$o_1 = (x_1^2 y_2 - x_1^2 y_3 - x_2^2 y_1 + x_2^2 y_3 + x_3^2 y_1 - x_3^2 y_2 + y_1^2 y_2 - y_1^2 y_3 - y_1 y_2^2 + y_1 y_3^2 + y_2^2 y_3 - y_2 y_3^2) / (2x_1 y_2 - 2x_1 y_3 - 2x_2 y_1 + 2x_2 y_3 + 2x_3 y_1 - 2x_3 y_2) \quad (1)$$

$$o_2 = (-x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 - x_1 x_3^2 + x_1 y_2^2 - x_1 y_3^2 - x_2^2 y_3 + x_2 x_3^2 - x_2 y_1^2 + x_2 y_3^2 + x_3 y_1^2 - x_3 y_2^2) / (2x_1 y_2 - 2x_2 y_1 - 2x_1 y_3 + 2x_3 y_1 + 2x_2 y_3 - 2x_3 y_2) \quad (2)$$

$$r = \sqrt{(x_1 - o_1)^2 + (y_1 - o_2)^2} \quad (3)$$

之后迭代计算每个轮廓点与假定圆心和半径的差异，保留符合圆形几何特性的点。通过设定阈值，比较剩余点在轮廓中的占比，得到最终的圆参数。

在圆和圆弧的区分上，使用矩形模板进行检测。模板的中心与圆心的距离保持不变，通过扫描图像来判断圆的完整性，最终识别结果如图 11 所示。

4. 基于卷积神经网络的字符识别

4.1. 数据集的建立

人工标注耗时且难以满足数据量要求，因此数据集中加入了来自 CVSSP 数据库的手写和标准字符，对于需要手动标注的数据集，将字符串切分为单个字符图像[20]。采用垂直投影法，如图 12，通过计算每列像素数生成投影直方图，并寻找空白部分作为字符间的切分间隔，之后统一格式及像素后分类保存，确保特征能够被最顶层的特征检测子感知。



Figure 12. Segmentation of characters and images
图 12. 字符图像切分

4.2. 字符识别模型

卷积神经网络[21]是当下最流行的深度神经网络之一，如图 13，其由卷积层、池化层和全连接层组成。

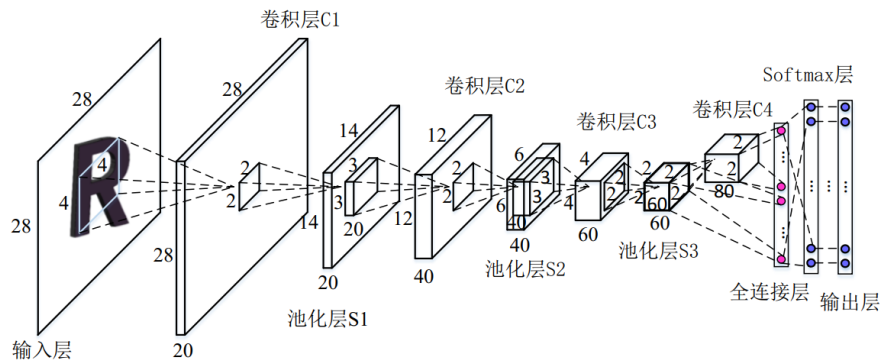


Figure 13. Convolutional neural network structure diagram
图 13. 卷积神经网络结构示意图

卷积层通过对前一层的特征映射进行卷积运算处理[22]，并与非线性活化功能结合，生成输出特征映射。以下是其运作步骤：

$$x_j^L = f \left(\sum_{i \in M_j} x_i^{L-1} * k_{ij}^L + b_j^L \right) \quad (4)$$

式中 f 为激活函数。“ $*$ ”是卷积符号， M_j 是输入的特征图子集， b_j^L 是卷积后特征图的偏置， k_{ij}^L 是卷积核矩阵。

通常情况下，CNN 架构中的池化层是在卷积层之后接入的。池化层的作用是通过下操作降低特征图的空间维度，从而减少参数数量、计算复杂程度、增加模型的平移不变性。在池化层中，通常进行求解平均值或极值的操作，具体如下：

$$x_j^L = f \left(\beta_j^L d \left(x_i^{L-1} \right) + b_j^L \right) \quad (5)$$

式中 β 是下采样层的权重系数， x_i^{L-1} 是输入的特征图， b_j^L 是卷积后特征图的偏置， d 是下采样函数。

卷积层提取的特征将被池化层进一步下采样，最后传输到全连接层，确定最终输出，如下式所示，式中 ω^L 是全连接层的权重系数。

$$x^L = f \left(\omega^L x^{L-1} + b^L \right) \quad (6)$$

训练字符识别模型的过程分为正向和反向两个阶段[23]，在正向传播阶段，为了便于处理，假设一个字符 (P, Q_n) ，其中 P 为输入， Q_n 为目标输出。计算 P 的实际输出 R_n 。 P 传入到正向网络中，每当训练完一层，就将该层的权值矩阵相乘，全部完成后，将参数传到输出层得到 R_n ，该流程依据的迭代公式如下：

$$R_n = f_n \left(\cdots \left(f_2 \left(f_1 \left(P_n \omega^1 \right) \omega^2 \right) \cdots \right) \omega^m \right) \quad (7)$$

Table 1. Diagram of the convolution layer and the pooling layer

表 1. 卷积层和池化层的相关数据表

网络结构层	滤波器大小	滤波器个数	连接数
C1	2×2	20	78,400
S1	3×3	20	39,200
C2	2×2	40	28,800
S2	3×3	40	14,400
C3	2×2	60	4800
S3	2×2	60	1200
C4	2×2	80	1600

接下来就是反向传播阶段，该阶段主要计算训练后的 R_n 与目标输出 Q_n 之间的偏差。根据极小误差法对参数矩阵进行反向调节，得到最终参数，当网络训练相对稳定时就结束训练，其中卷积层和池化层中的滤波器大小、连接数等数据如表 1 所示。

随着训练周期的增加，该模型的权重更新次数也会增加，所以可能会出现过拟合的现象。该模型训练过程中训练周期、迭代次数、准确率等之间的关系如表 2 所示。可知，当训练周期达到 18 次时，准确率接近峰值，18 次之后有所下降，所以最终将模型的训练周期设置为 18，结果达到 98.65% 的准确率。

用训练好的模型来识别字符，输入原图像以轴类工程图纸为例，由于该类零件的图纸信息不复杂，所以可以直接进行字符识别，识别结果如图 14 所示。

Table 2. Table of relationships between training cycles, iterations, accuracy, etc.

表 2. 训练周期、迭代次数和准确率等之间的关系表

周期	迭代次数	识别准确率	损失误差	训练时间	学习率
1	250	7.62%	3.4831	0:01:04	0.01
2	550	41.58%	2.7820	0:02:23	0.01
4	1100	87.14%	0.5388	0:03:44	0.01
6	1650	89.95%	0.3671	0:05:16	0.01
8	2200	91.43%	0.2833	0:09:42	0.01
10	2800	93.72%	0.1542	0:11:12	0.01
12	3350	96.09%	0.0652	0:13:04	0.01
14	3900	98.23%	0.0307	0:16:89	0.01
16	4450	98.44%	0.0542	0:19:73	0.01
18	5050	98.65%	0.0173	0:22:55	0.01
20	5620	98.44%	0.01406	0:24:66	0.01

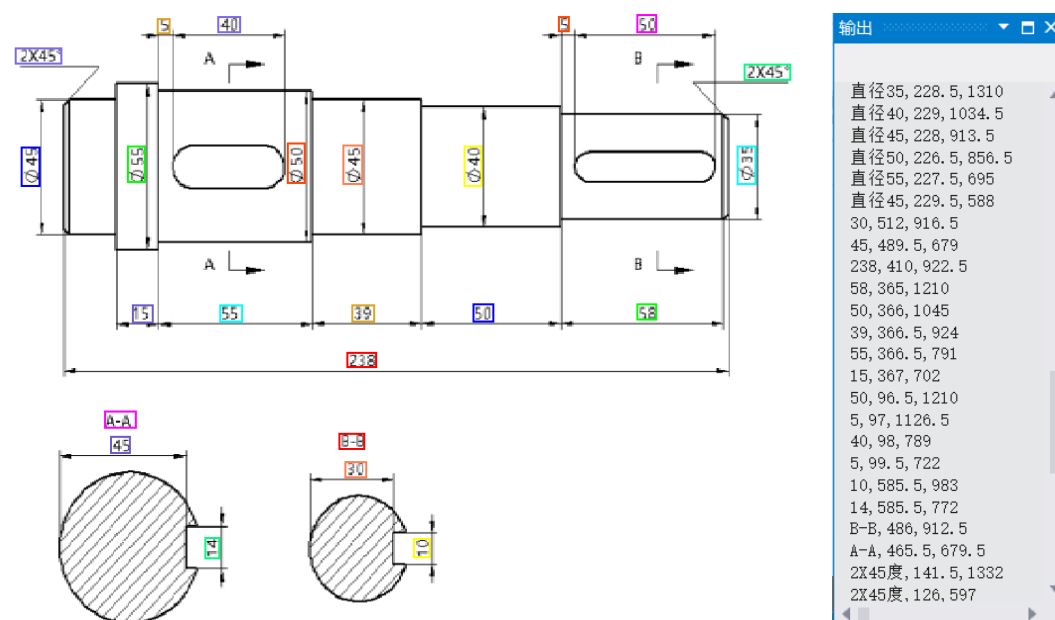


Figure 14. Character recognition result

图 14. 字符识别结果

5. 知识引导下的三维重建

5.1. 工程语义

在二维图纸的投影关系中,零件的主要平面与投影面平行,主要轴线与投影面垂直,从而导致投影的重合,使得大量信息积聚。而且基元间的组合也具有多种形式,使得投影产生模糊性。此外,工程图纸中的尺寸标注只是对直线、圆、圆弧的约束,隐式地表达了形体的大小,但尺寸本身的三维特性、功能语义特性是隐含的。从图纸的表达方式来讲,它是以视图为单位的,一般是基本三视图,复杂的模型还需要剖视图、局部视图、轴测图等。对于计算机来说,这些视图是零散的,并没有直接的联系,通过语义的理解,它们才能构成图纸的基本图元,这些表达方式的相关特征如表3所示。

显示尺寸语义即尺寸本身表达的语义。尺寸数字通过前缀与后缀可以将其约束的基元体进行分类，如表 4 所示。对于隐式尺寸的语义知识，主要通过尺寸与图形之间关联来获得，工程图中的尺寸标注是用来描绘三维形体的外形和特征间拓扑关系的，例如当复合特征标注如“Φ6 沉孔 Φ12 深 4”，通过识别到的字符文本可以搜索到相应的结构尺寸。

Table 3. Related features of different expressions
表 3. 不同表达方式的相关特征

视图类型	投影方式	剖视占比	适用对象	视图特征	应用条件
基本视图	直接投影	0	表达外形	无剖面线	无需表达内形
剖视图	剖后再投影	50 或 100	表达内形或表达内外形	剖面线起止轮廓线	无需表达外形
局部放大图	视图或剖视图，与放大部位的表达无关	0 或 100	表达细节	罗马数字与波浪线分界	需放大结构
断裂视图	剖后不投影	100	表达断面	波浪线分界	无需表达其余结构
向视图、斜视图	相应处理	0	表达外形	无剖面线	无须表达外形

Table 4. Explicit size semantics
表 4. 显式尺寸语义

尺寸分类	前缀	后缀	对应的特征类型
直径	Φ	无	圆柱、圆锥、环体
半径	R	无	圆柱
线性	无	无	平面立体
角度	无	°	斜面
倒角	C	无	圆锥体
螺纹	M、Tr、SG、ZG	无	螺纹

扫描图纸识别后的图素都是孤立存在的，为了能够实现二维理解，需要建立图素之间的拓扑关系，图 15 汇总了图素之间的拓扑关系。

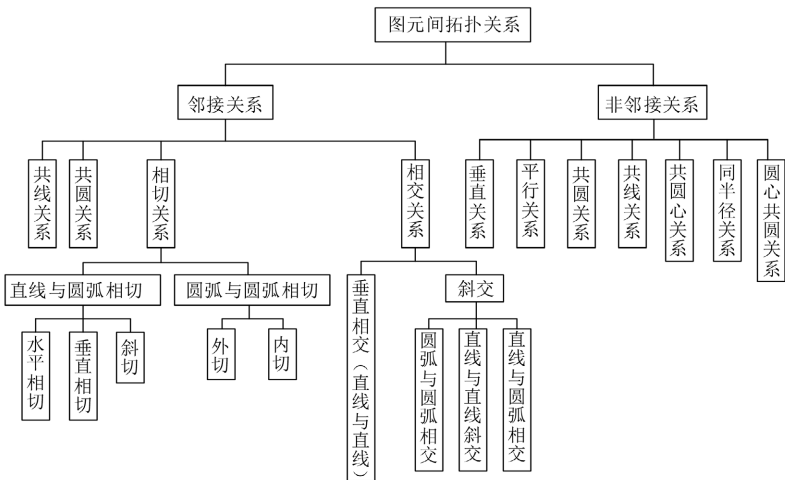


Figure 15. The topological relationship between primitives
图 15. 图元间的拓扑关系

5.2. 基元重建

在工程图中，平面投影根据正投影规则可分为三类，如图 16。图(a)和图(b)中的投影平面垂直于坐标面，图(c)中的投影平面则不垂直于任何坐标面。垂直坐标面的投影平面在三视图中呈现直线，而不垂直坐标面的投影平面则呈现三个封闭环。

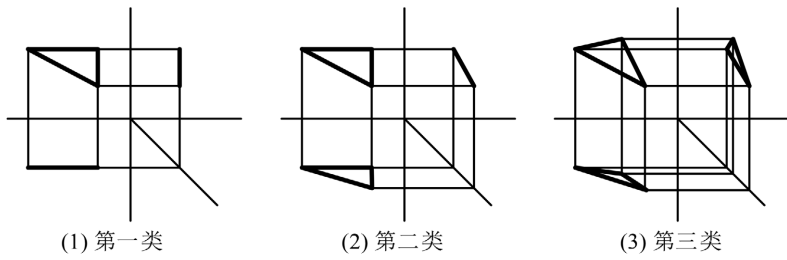


Figure 16. Projection view of the plane
图 16. 平面的投影视图

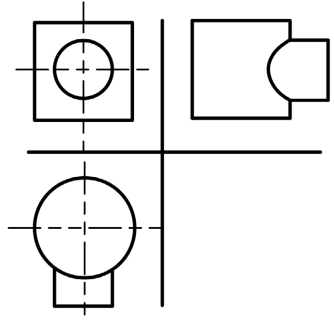


Figure 17. Cylindrical intersecting
图 17. 圆柱相贯

对于基本平面，投影特性可以从含有直线的视图出发，在其他视图中搜索对应的封闭环，实现简单基元的重建。然而，大多数平面并非完整投影，且更多的是组合基元，重建时不仅需要投影约束关系，还需要结合其他语义信息辅助识别。

如图 17，圆柱面与平面截交时，矩形投影可能不完整，但仍是封闭环，且由直线构成，圆的投影可能变成圆弧。两圆柱面相贯时，矩形投影形成封闭环，包含部分圆弧线段[24]。

由于直接从组合的基元进行重建十分困难，所以先识别单个圆柱面投影的图素，重建出单个圆柱，再识别相贯线或者截交线，重建对应的几何特征，最后用简单的布尔运算获得组合基元的三维模型。

6. 基于轴类产品的开发与实现

如图 18，系统分为四个模块：图纸输入模块、图纸预处理模块、图纸识别模块、三维重建模块。整个三维重建方法是基于精确的二维重建实现的，所以对于扫描图纸的识别是重点，原扫描图纸如图 19 所示。

利用机械中的知识，首先对二维图形对应的三维形体进行假设，然后搜索扫描图纸的字符识别结果，基于实例和规则进行推理，匹配相应的特征，得到的三维特征数据如图 20 所示。

特征匹配完成后，按指定的设计规则进行特征的重建，同时特征之间的拓扑关系在二维重建过程中已经得到，故三维重建时可以直接根据上述识别结果进行重建，对于特征的凹凸性判别，由于键槽特征始终是凹特征，所以直接进行拉伸切除就可以实现，最终的三维重建模型如图 21 所示。



Figure 18. 3D reconstruction of the system interface

图 18. 三维重建系统界面

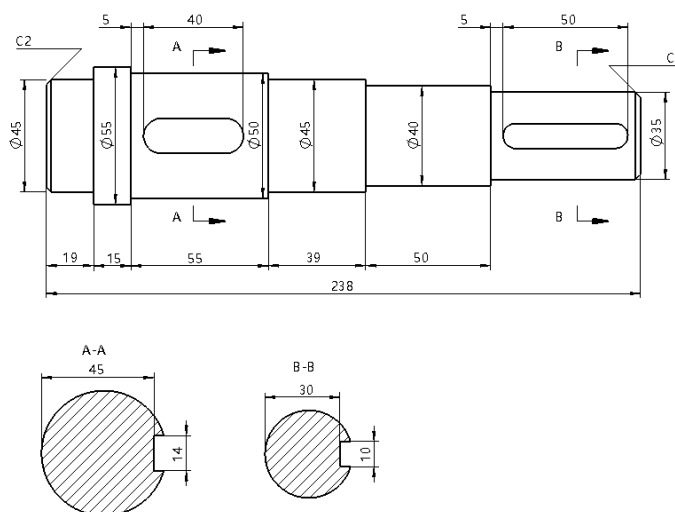


Figure 19. Scan drawing example 1

图 19. 扫描图纸实例 1

特征	直径	长度/角度	距离	高度	父特征
圆柱1	45	19	null	null	null
圆柱2	55	15	null	null	null
圆柱3	50	55	null	null	null
圆柱4	45	39	null	null	null
圆柱5	40	50	null	null	null
圆柱6	35	60	null	null	null
键槽1	14	40	5	5	圆柱3
键槽2	10	50	5	5	圆柱6
倒角1	2	45°	null	null	圆柱1
倒角2	2	45°	null	null	圆柱6

Figure 20. Result of example 1 identification

图 20. 实例 1 识别结果

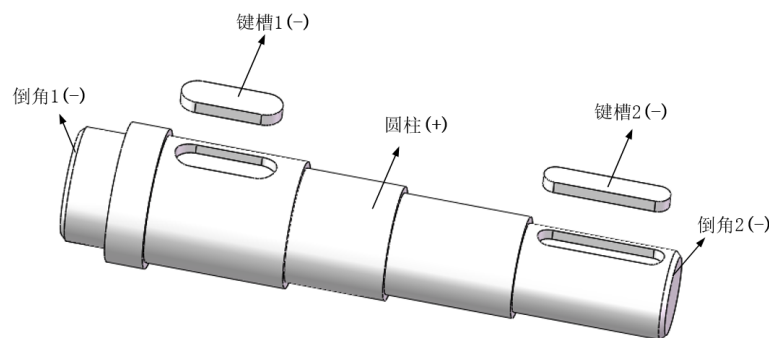


Figure 21. Example 1 3D reconstruction diagram
图 21. 实例 1 三维重建示意图

7. 结论

在机械知识的引导下,以扫描图纸为系统输入,利用图像识别技术完成扫描图纸的高层理解,最后基于轴类实例与规则推理实现三维重建。研究结论如下:

- 1) 采用基于 Hough 变换原理的改进方法,解决了传统 Hough 变换对噪声敏感的问题,并且利用图形的几何特性,降低算法复杂度,大大提高了识别效率与准确度。
- 2) 提出了符合工图特殊情景的基于卷积神经网络的字符识别方法,结果表明该模型可以对简单工图直接进行识别字符,能够有效的兼容干扰图线,识别率达 98%左右。
- 3) 结合工程语义知识,将基元按特征分类并给出不同重建策略,最终实现三维重建。

参考文献

- [1] Kapitanov, A., Kozlova, A. and Tyasto, S. (2018) The Information Model of the Modern Digital Production. *MATEC Web of Conferences*, **224**, Article 02065. <https://doi.org/10.1051/mateconf/201822402065>
- [2] 杨建鸣, 陆家山, 严鹏贺. 工程图的三维重建方法研究[J]. 机械设计与制造, 2017(12): 216-219.
- [3] Idesawa, M. (1973) A System to Generate a Solid Figure from Three View. *Bulletin of JSME*, **16**, 216-225. <https://doi.org/10.1299/jsme1958.16.216>
- [4] Wesley, M.A. and Markowsky, G. (1981) Fleshing Out Projections. *IBM Journal of Research and Development*, **25**, 934-954. <https://doi.org/10.1147/rd.256.0934>
- [5] 刘世霞, 胡事民, 汪国平, 孙家广. 基于三视图的三维形体重建技术[J]. 计算机学报, 2000, 23(2): 141-146.
- [6] 龚洁晖, 张慧. 基于工程图的三维重建研究[J]. 软件学报, 2008, 19(7): 1794-1805.
- [7] Alexei, Z. and Arkady, Z. (2014) Three-Dimensional Reconstruction from Projections Based on Incidence Matrices of Patterns. *AASRI Procedia*, **9**, 72-77. <https://doi.org/10.1016/j.aasri.2014.09.013>
- [8] Aldefeld, B. (1983) On Automatic Recognition of 3D Structures from 2D Representations. *Computer-Aided Design*, **15**, 59-64. [https://doi.org/10.1016/0010-4485\(83\)90169-0](https://doi.org/10.1016/0010-4485(83)90169-0)
- [9] 张雪, 赵德胜, 李丽萍. 基于 CSG 方法的工程图三维形体重建[J]. 西安邮电大学学报, 2015, 20(4): 80-83, 124.
- [10] Chen, Z. and Perng, D. (1988) Automatic Reconstruction of 3D Solid Objects from 2D Orthographic Views. *Pattern Recognition*, **21**, 439-449. [https://doi.org/10.1016/0031-3203\(88\)90003-9](https://doi.org/10.1016/0031-3203(88)90003-9)
- [11] Lin, W. and Fang, C. (2007) Synthesized Affine Invariant Function for 2D Shape Recognition. *Pattern Recognition*, **40**, 1921-1928. <https://doi.org/10.1016/j.patcog.2006.03.021>
- [12] Chen, G., Zhang, D., Li, P., Chen, C., Wang, Y. and Su, S. (2019) Research on the Knowledge Representation and Retrieval for Mechanical Product Design Based on Ontology. *International Journal of Wireless and Mobile Computing*, **16**, 340. <https://doi.org/10.1504/ijwmc.2019.10021628>
- [13] 杨海成, 主编. 数字化设计制造技术基础[M]. 西安: 西北工业大学出版社, 2009.
- [14] Suresh, K. and Srinivasa rao, P. (2018) Various Image Segmentation Algorithms: A Survey. In: Satapathy, S., Bhateja, V. and Das, S., Eds., *Smart Intelligent Computing and Applications*, Springer, 233-239.

- https://doi.org/10.1007/978-981-13-1927-3_24
- [15] Karthikeyan, V. (2013) Hilditch's Algorithm Based Tamil Character Recognition. *International Journal of Computer Science & Engineering Technology*, **4**, 268-273.
- [16] Liu, Z., Liang, J. and Pan, Y. (2014) Construction of Thermodynamic Properties Look-Up Table with Block-Structured Adaptive Mesh Refinement Method. *Journal of Thermophysics and Heat Transfer*, **28**, 50-58.
<https://doi.org/10.2514/1.t4273>
- [17] Oh, J.S. (2018) Straight Line Detection Using PCA and Hough Transform. *Journal of the Korea Institute of Information and Communication Engineering*, **22**, 227-232.
- [18] Jiang, G., Wang, X., Wang, Z. and Liu, H. (2016) Wheat Rows Detection at the Early Growth Stage Based on Hough Transform and Vanishing Point. *Computers and Electronics in Agriculture*, **123**, 211-223.
<https://doi.org/10.1016/j.compag.2016.02.002>
- [19] Djekoune, A.O., Messaoudi, K. and Amara, K. (2017) Incremental Circle Hough Transform: An Improved Method for Circle Detection. *Optik*, **133**, 17-31. <https://doi.org/10.1016/j.ijleo.2016.12.064>
- [20] Sagar, S. and Dixit, S. (2019) A Comprehensive Study on Character Segmentation. In: Pandian, D., Fernando, X., Baig, Z. and Shi, F., Eds., *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*, Springer International Publishing, 1509-1515.
https://doi.org/10.1007/978-3-030-00665-5_141
- [21] Li, P., Peng, L. and Wen, J. (2016) Rejecting Character Recognition Errors Using CNN Based Confidence Estimation. *Chinese Journal of Electronics*, **25**, 520-526. <https://doi.org/10.1049/cje.2016.05.018>
- [22] Ghiasi-Shirazi, K. (2019) Generalizing the Convolution Operator in Convolutional Neural Networks. *Neural Processing Letters*, **50**, 2627-2646. <https://doi.org/10.1007/s11063-019-10043-7>
- [23] Jiang, Z., Yin, F., Ge, Y. and Ren, L. (2020) A Multi-Scale Framework with Unsupervised Joint Training of Convolutional Neural Networks for Pulmonary Deformable Image Registration. *Physics in Medicine & Biology*, **65**, Article 015011.
<https://doi.org/10.1088/1361-6560/ab5da0>
- [24] 畅通, 包宏. 基于图形语义理解的三维重建技术[J]. 计算机工程与设计, 2005(9):2274-2276.