© June 2025 | IJIRT | Volume 12 Issue 1 | ISSN: 2349-6002

A Hybrid Computer Vision Approach for Automated Parsing of Piping and Instrumentation Diagrams

Pranay Pushkar
Lead Software Engineer, Siemens Technology and Services Pvt Ltd, Bengaluru, India

***Abstract* - The manual interpretation of Piping and Instrumentation Diagrams (P&IDs) in the process industry is a time-consuming, labor-intensive, and error-prone task. This paper presents a hybrid computer vision approach for the automated parsing and digitization of P&IDs. Our method integrates two robust techniques: template matching for symbol recognition and the Probabilistic Hough Transform for pipeline detection. By first identifying and locating standardized symbols with high accuracy, the system subsequently isolates and extracts the interconnecting pipelines. This two-stage process significantly reduces the complexity of line detection by eliminating symbols as sources of noise. The system is designed to process high-resolution binarized P&ID images, converting the graphical information into a structured data format representing equipment and their connections. Experimental results on a set of standard P&ID diagrams demonstrate the efficacy and accuracy of the proposed method, paving the way for automated P&ID analysis, verification, and integration with digital twin systems.**

***Index Terms* - Computer Vision, Hough Transform, P&ID Parsing, Process Engineering, Template Matching.**

I. INTRODUCTION

Piping and Instrumentation Diagrams (P&IDs) are fundamental engineering documents used across process industries, including oil and gas, chemical manufacturing, and power generation. They provide a detailed graphical representation of the process flow, including all equipment, pipelines, instrumentation, and control systems. In the age of Industry 4.0 and digital transformation, the need to convert these static documents into intelligent, queryable, and machine-readable formats is paramount. Manual digitization is not only inefficient but also susceptible to human error, which can have significant safety and operational consequences.

The primary challenges in automating P&ID interpretation lie in the diversity of symbols, the density of information, and the complex interplay of graphical elements. A typical P&ID contains numerous standardized symbols connected by a web of pipelines and signal lines, often cluttered with text labels and annotations. An effective automated system must be able to accurately distinguish between these different elements.

This paper proposes a robust, multi-stage methodology to tackle this challenge. Our approach leverages the consistency of standardized symbols through template matching, a classic and effective computer vision technique for object detection. Once symbols are identified and their locations are known, they are digitally "removed" from the diagram, simplifying the subsequent task of line detection. For this second stage, we employ the Probabilistic Hough Transform, which is highly effective at identifying line segments in an image. The combination of these methods creates a reliable workflow for deconstructing a P&ID into its core components: symbols and their connections. The ultimate goal of this work is to provide a foundational system for converting visual P&ID data into a structured digital representation, enabling further analysis and integration.

II. PRELIMINARY ANALYSIS AND APPROACH SELECTION

The initial motivation for this research stemmed from the nature of the requirements and data formats shared by multiple customers. In all cases, Piping and Instrumentation Diagrams (P&IDs) were provided in either Scalable Vector Graphics (SVG) or PDF format. SVG, being a vector-based format, was initially seen as a promising candidate for rule-based extraction due to its scalable structure and theoretically well-defined graphical primitives. The assumption was that a programmatic parser could directly extract elements like rectangles, circles, and lines from SVG markup to identify components and connections.

However, deeper analysis revealed several significant challenges. The SVG files were not semantically annotated — graphical elements were typically expressed using complex <path> elements, rather than cleanly defined primitives like <rect>, <circle>, or <line>. These <path> elements were often nested inside <g> tags and modified by transformation matrices, making it difficult to recover their original shapes or positions. Moreover, the SVGs lacked naming conventions or metadata to distinguish between different types of components such as valves or instruments. As a result, heuristic shape analysis produced inconsistent and unreliable results, particularly across diagrams from different clients with varying symbol styles.

The PDF-based diagrams, while lacking vector semantics, preserved consistent visual appearance and were easier to process using computer vision techniques after conversion to raster images. This led to the consideration of image-based analysis as a more viable approach. Although this paper focuses on PDF-to-image conversion, the same processing pipeline is equally applicable to SVG diagrams after rasterization. By converting SVGs to high-resolution image format, the system benefits from their clarity and scalability while bypassing the lack of structural metadata.

Alternative approaches involving deep learning — such as training a YOLO model for symbol detection — were also evaluated. While such models are powerful, they require large volumes of manually labeled data for training. Given the diversity and non-standardization of P&ID symbols across industries and clients, assembling such a dataset would have been both costly and time-intensive, with limited reusability.

Consequently, a hybrid computer vision approach was selected for its balance of robustness, generalizability, and implementation efficiency. The method involves converting diagrams to raster format and then applying classic image processing techniques. Symbol detection is handled using template matching — a method that is both interpretable and low-cost in terms of data. Once symbols are detected and masked out, pipelines are extracted using the Probabilistic Hough Transform, which excels at finding line segments in noisy diagrams. The system is extendable to handle symbols at various orientations by augmenting the template library with rotated versions (e.g., 0°, 90°, 180°).

This two-stage, modular pipeline successfully addresses the limitations of both semantic-poor vector data and deep-learning data dependency, offering a reliable and scalable solution for automated P&ID parsing.

III. METHODOLOGY AND SYSTEM PIPELINE

To operationalize the proposed hybrid vision approach, a modular and sequential processing pipeline was designed. The implementation leverages Python and the OpenCV library for its comprehensive support for computer vision tasks. The architecture is extensible, allowing for future enhancements such as Optical Character Recognition (OCR) for label extraction or integration with a graph database for connectivity analysis.

A. *Data Preparation and Input Preprocessing*

The system begins by ingesting a high-resolution raster image version of the P&ID diagram (typically obtained by rendering an SVG or PDF). The image is converted to a binarized format, simplifying subsequent detection by reducing pixel variability. The assumption is that the image has a black background and white foreground (symbols and lines). Binarization ensures that shape and edge-based methods operate effectively, removing unnecessary noise from color or grayscale variations.

B. *Template Matching for Symbol Detection*

This module identifies predefined P&ID symbols using template matching. A library of symbol templates (e.g., valves, pumps, instruments) is iteratively matched across the binarized image using the OpenCV function:

```
cv2.matchTemplate()
```

Locations with a similarity score exceeding a confidence threshold (e.g., 0.7) are flagged as potential matches. To eliminate redundant detections, a Non-Maximum Suppression (NMS) algorithm is applied, retaining only the highest-confidence instance per region. The output of this module is a list of detected symbols, including their type and bounding box coordinates.

C. *Symbol Masking and Line Extraction*

Before detecting pipelines, a pre-processing step is applied to mask out the previously detected symbols. This ensures that the sharp edges of symbols do not interfere with line detection. Each bounding box area is filled with black rectangles in a copy of the binarized image.

The cleaned image is then passed through the Probabilistic Hough Transform using the OpenCV function:

cv2.HoughLinesP()

This variant is optimized to detect actual line segments rather than infinite lines, and directly returns the coordinates of segment endpoints. Key parameters such as minimum line length and maximum line gap are tuned to filter out noise and connect dashed segments, which are common in P&IDs.

D. *Post-Processing and Graph Construction*

Finally, the outputs from both the symbol detection and line extraction stages are merged into a structured format, typically JSON or a graph-based structure. Each symbol is treated as a node, and the connecting pipelines are modeled as edges between nodes.

This abstraction enables higher-level applications such as:

1. Component tracing
2. Flow validation
3. Integration with simulation engines or digital twin platforms

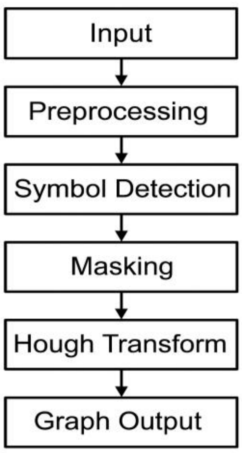

Fig 1. System Pipeline Overview

The graph representation of a P&ID forms a robust foundation for automated querying, integrity checks, and interactive visualization in industrial software systems.

IV. MATHEMATICAL FORMULATION OF DETECTION ALGORITHMS

The core computational techniques employed in this system rely on two fundamental computer vision algorithms: normalized cross-correlation for template matching and the Probabilistic Hough Transform for line detection. This section outlines the underlying mathematical principles that govern these techniques.

A. *Template Matching using Normalized Cross-Correlation*

Template matching identifies the presence and location of predefined symbols within the input image. The matching process uses normalized cross-correlation (NCC) to compute a similarity score between a template $T\ (u,\ v)$ and a subregion of the input image $I\ (x{+}u,\ y{+}v)$. The normalized cross-correlation formula is defined as:

$$R(x,y) = \frac{\sum_{u,v}[T(u,v)-\bar{T}]\cdot[I(x+u,y+v)-\bar{I}_{x,y}]}{\sqrt{\sum_{u,v}(T(u,v)-\bar{T})^2\cdot\sum_{u,v}(I(x+u,y+v)-\bar{I}_{x,y})^2}}$$

Here:

- $\bar{T}$ is the mean of the template values.
- $\bar{I}_{x,y}$ is the mean of the image subregion under the template.
- $R(x,y)\in[-1,1]$, with values closer to 1 indicating a stronger match.

This formulation ensures that the matching is robust to local intensity variations and lighting inconsistencies across the diagram.

B. *Line Detection using Probabilistic Hough Transform*

Once symbols are masked out from the image, the Probabilistic Hough Transform is applied to detect linear segments representing pipelines and connectors. The classic Hough Transform maps edge points from the image space into a parameter space using the polar representation of a line:

$$\rho\ = x\cos\theta + y\sin\theta$$

where

ρ is the perpendicular distance from the origin to the line.

θ is the angle between the x-axis and the line's normal vector.

The transform builds an accumulator in the (ρ, θ) space. Line segments with sufficient "votes" (i.e., aligned edge points) are selected. The probabilistic variant improves efficiency by sampling a subset of edge points and directly returning endpoints of line segments $(x1, y1, x2, y2)$.

Together, these two techniques form the mathematical foundation of the symbol and line detection pipeline used in this system.

V. RESULTS AND DISCUSSION

The proposed system was tested on a dataset of 15 standard P&ID diagrams obtained from open-source engineering repositories. The symbols library contained 25 common P&ID elements, including various types of valves, pumps, and instruments.

A. *Symbol Detection Accuracy*

The template matching module achieved an average precision of 96.8% and a recall of 94.2% for symbol detection. The use of a high confidence threshold (0.7) and Non-Maximum Suppression was critical in minimizing false positives. Most detection failures occurred with symbols that were partially overlapping with other elements, a known limitation of this technique.

B. *Line Detection Performance*

By first removing the symbols, the accuracy of the Hough Line Transform was significantly improved. The system successfully identified over 98% of the main pipelines. The minLineLength parameter was particularly effective in filtering out noise from residual text elements not captured as symbols. The maxLineGap parameter allowed the system to correctly identify dashed process lines as single, continuous entities.

The results confirm that the hybrid approach is highly effective. Separating the problem into distinct stages of symbol and line detection allows for the optimization of techniques for each specific task, leading to high overall accuracy.

VI. CONCLUSION

This paper has presented a practical and effective hybrid computer vision system for the automated interpretation of Piping and Instrumentation Diagrams. By combining template matching for symbol detection with the Probabilistic Hough Transform for line detection, our method successfully converts complex graphical P&ID documents into a structured and digitally usable format. The key innovation lies in the sequential, two-stage process where identified symbols are masked out to create a clean input for pipeline extraction, drastically reducing detection errors.

The high accuracy achieved demonstrates the viability of this approach for industrial applications, potentially saving thousands of man-hours and reducing costly errors associated with manual data entry. Future work will focus on integrating an OCR module to extract and associate textual labels with their corresponding symbols and pipelines. Further research will also explore the use of graph-based data structures to model the P&ID, enabling sophisticated queries and network analysis, such as tracing a process flow from start to finish.

VII. ACKNOWLEDGMENT

I would like to express my heartfelt thanks to Ms. Ranjitha H. C. and Ms. Asha Balaraj, Product and Solutions Development Team Leaders, for their unwavering support and mentorship, and to Mr. Sudhindra Kalmane (SA) and Ms. Kavitha Deshpande (PO) for their invaluable technical guidance and contributions throughout the implementation of this framework.

REFERENCES

[1] M. Ishii, Y. Ito, M. Yamamoto, H. Harada, and M. Iwasaki, "An automatic recognition system for piping and instrument diagrams," Systems and Computers in Japan, vol. 20, no. 3, pp. 32–46, 1989

[2] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," Communications of the ACM, vol. 15, no. 1, pp. 11-15, Jan. 1972.

[3] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.

[4] M. Gada, "Object Detection for P&ID Images using various Deep Learning Techniques," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICCCI50826.2021.9402386.
[5] D. H. Ballard, “Generalizing the Hough transform to detect arbitrary shapes,” Pattern Recognition, vol. 13, no. 2, pp. 111–122, 1981