

SIMC2024 Report: A Heuristic Approach to Pattern Classification and Recovery

Li Shiming
xajdfz

Liu Ye
xajdfz

Wang Boran
xajdfz

Abstract

This paper presents a unified, heuristic approach to solving image pattern challenges. We treat the problems not as complex mathematical abstractions, but as data transformation tasks. For Task 1 (Matrix) and Task 2 (Vector), we employ a direct "Key-Matching" strategy by geometrically transforming a reference image. For the unsupervised Task 3, we adopt a "Data Detective" approach: determining the image topology from vector length, applying K-Means clustering to reveal hidden structures, and using signal averaging to recover a latent image from noise.

1 Introduction

We are presented with a dataset of patterns that have been rotated and mixed. The core challenge is to identify the orientation ($0^\circ, 90^\circ, 180^\circ, 270^\circ$) under different data conditions:

1. **Task 1:** The data is clean and provided as 2D images.
2. **Task 2:** The data is flattened into 1D vectors, hiding the spatial structure.
3. **Task 3:** The data is flattened, noisy, and unlabeled (we don't know which image is which).

2 Methodology

2.1 Task 1: The "Master Key" Strategy

Instead of training a complex classifier, we observe that the rotations are discrete. We simply take the first image in the dataset and treat it as our "Master Key" (Reference). By manually rotating this Master Key by 90, 180, and 270 degrees, we generate a Reference Dictionary containing 4 perfect templates. We then classify every other image in the dataset by checking which of the 4 templates it matches perfectly.

2.2 Task 2: The "Fold and Unfold" Trick

Task 2 complicates matters by flattening the $N \times N$ image into a long vector of length $L = 1296$.

You cannot "rotate" a 1D line of numbers. Our solution is topological:

1. Fold (Reshape): We calculate $N = \sqrt{1296} = 36$. We fold the 1D vector back into a

36×36 matrix.

2. Rotate: Now that structure is restored, we rotate the matrix using the Task 1 method.
3. Unfold (Flatten): We flatten the rotated matrix back into a 1D vector to compare it with the dataset.

This "Reshape-Rotate-Flatten" cycle allows us to reuse the efficient matching logic from Task 1.

2.3 Task 3: Unsupervised Latent Pattern Discovery

Task 3 is the most challenging: the data is noisy, flattened ($L = 1089$), and we have no "Master Key". We deduce the dimensions are 33×33 ($33^2 = 1089$). Since we cannot match against a known reference, we group the data based on similarity.

We use K-Means Clustering ($K = 4$) to automatically separate the data into 4 distinct groups. We hypothesize that the Centroid (average) of each cluster will represent the "clean" version of that rotation, as averaging cancels out the random noise.

3 Algorithm Description

The logic for our solution is streamlined into two core algorithms.

3.1 Supervised Matching (Task 1 & 2)

Algorithm 1 Generate and Match

```
1: Input: Dataset  $P$ , First Image  $Ref$ 
2: Output: Counts for each angle
3:  $Templates \leftarrow \{\}$ 
4: for  $angle$  in  $[0, 90, 180, 270]$  do
5:    $Img \leftarrow \text{Rotate}(Ref, angle)$ 
6:   if Data is Flat then
7:      $Img \leftarrow \text{Flatten}(Img)$                                  $\triangleright$  The "Unfold" step
8:   end if
9:    $Templates[angle] \leftarrow Img$ 
10: end for
11: for  $p$  in  $P$  do
12:   Find  $angle$  where  $p == Templates[angle]$ 
13:   Increment count for  $angle$ 
14: end for
```

3.2 Unsupervised Recovery (Task 3)

Algorithm 2 Cluster and Recover

- 1: **Input:** Noisy Matrix \mathbf{X}
 - 2: $Labels \leftarrow \text{KMeans}(\mathbf{X}, K = 4)$
 - 3: ▷ Sort data to visualize structure
 - 4: $\mathbf{X}_{sorted} \leftarrow \text{SortByLabels}(\mathbf{X}, Labels)$
 - 5: ▷ Recover signal by averaging
 - 6: $CleanVector \leftarrow \text{GetCentroid}(\text{Cluster}_0)$
 - 7: $RecoveredImage \leftarrow \text{Reshape}(CleanVector, (33, 33))$
-

4 Implementation and Results

4.1 Tasks 1 & 2: Exact Matching

For the 2D matrices (Task 1), our Master Key strategy achieved 100% matching accuracy instantly. For the flattened vectors (Task 2), the "Fold and Unfold" strategy proved robust. By temporarily restoring the 2D geometry, we successfully bridged the gap between vector storage and spatial transformation.

4.2 Task 3: From Chaos to Order

In Task 3, we visualized the dataset as a "Design Matrix". Initially, it appeared as random noise (high entropy). After applying K-Means and sorting the matrix by cluster labels, distinct bands appeared (low entropy), visually proving the existence of 4 orientation classes.

Most importantly, the Cluster Centroid reconstruction was highly successful. While individual samples were noisy, the average of the cluster revealed a clear, distinct image (an animal figure). This confirms that simple statistical averaging is a powerful tool for signal recovery in unsupervised tasks.

5 Conclusion

We approached these tasks with a philosophy of simplicity. For Tasks 1 and 2, we showed that complex problems often have simple geometric solutions (rotation and reshaping). For Task 3, we demonstrated that even without labels, the inherent structure of data can be discovered through clustering. The "Design Matrix" visualization provided a powerful intuitive verification of our results.

6 AI Use Report

Trae IDE was used for inline code completions. **Gemini 3.0** was used to generate L^AT_EX code in formatting the document.