Computer Vision Assignment-2 Report

Sushovit Nanda

October 5, 2025

1 Introduction

This report presents the implementation and analysis of two computer vision tasks: (1) Integral Image computation and Haar-like feature extraction, and (2) Texture classification using various feature extraction methods on the **KTH-TIPS** dataset.

2 Question 1: Integral Image and Haar-like Features

2.1 Part 1: Integral Image Computation and Verification

The integral image was successfully implemented and verified for correctness. The implementation computes the cumulative sum of pixel values in both horizontal and vertical directions, enabling efficient computation of rectangular region sums in constant time.

Results:

- Test image dimensions: 256×256 pixels
- Verification performed on multiple regions with perfect accuracy
- Region (10,10) to (50,50): Integral = 68490.00, Direct = 68490.00, Difference = 0.000000
- Region (30,30) to (100,100): Integral = 511237.00, Direct = 511237.00, Difference = 0.000000
- Region (0,0) to (99,99): Integral = 773332.00, Direct = 773332.00, Difference = 0.000000

The integral image computation was verified to be mathematically correct with zero error tolerance, confirming the efficiency of the algorithm for rectangular region sum calculations.

2.2 Part 2: Haar-like Feature Extraction

Four different Haar-like feature patterns were implemented and extracted from the center 50×50 region of the test image:

1. **Horizontal pattern [W—B]:** White and black regions side by side
2. **Vertical pattern [W/B]:** White and black regions stacked vertically
3. **Three-rectangle pattern [W—B—W]:** Three horizontal regions with alternating intensities
4. **Four-rectangle pattern [W—B/B—W]:** 2×2 grid pattern with diagonal intensity arrangement

Results:

- Filter size: 24×24 pixels
- Center region: 50×50 pixels
- Total features extracted: 784 features
- Features per pattern: 196 features
- Feature value range: Min=-5248.00, Max=26368.00, Mean=5171.51

The Haar-like features successfully captured local intensity patterns and edge information from the center region of the image, demonstrating their effectiveness for feature representation.

3 Question 2: Texture Classification

The texture classification experiment was performed on the **KTH-TIPS** dataset containing 810 images from 10 texture classes: aluminium_foil, brown_bread, corduroy, cotton, cracker, linen, orange_peel, sandpaper, sponge, and styrofoam. All images were resized to 200×200 pixels for consistency.

3.1 Part 2(i): Raw Pixel Intensity Classification

Raw pixel values were used directly as features by flattening each 200×200 image into a 40,000-dimensional feature vector.

Results:

- Feature dimensions: 40,000 features per image
- Training samples: 567, Test samples: 243
- Classification accuracy: 42.39%

The raw pixel approach achieved the lowest accuracy, indicating that pixel intensities alone are insufficient for texture classification due to their sensitivity to illumination and lack of structural information.

3.2 Part 2(ii): Local Binary Pattern (LBP) Classification

LBP features were computed using 8 neighboring points with radius 1, creating 256-dimensional histogram features that capture local texture patterns.

Results:

- LBP feature dimensions: 256 features per image
- Classification accuracy: 93.83%

LBP achieved the highest accuracy among all methods, demonstrating its effectiveness for texture classification due to its illumination invariance and ability to capture local texture patterns.

3.3 Part 2(iii): Bag of Words Classification

Dense gradient features were extracted from image patches and clustered using K-means to create a visual vocabulary of 50 words, then represented as normalized histograms.

Results:

- Vocabulary size: 50 visual words
- BoW feature dimensions: 50 features per image
- Classification accuracy: 67.90%

The Bag of Words approach provided moderate performance, capturing global texture characteristics through local feature aggregation.

3.4 Part 2(iv): Histogram of Oriented Gradients (HOG) Classification

HOG features were computed using 9 orientation bins and 16×16 pixel cells, creating gradient orientation histograms that capture edge and shape information.

Results:

- HOG feature dimensions: 1,296 features per image
- Classification accuracy: 79.42%

HOG features provided good classification performance by capturing gradient orientation patterns, which are effective for texture discrimination.

3.5 Performance Comparison

Method	Accuracy	Feature Dimensions
Raw Pixels	42.39%	40,000
LBP	93.83%	256
Bag of Words	67.90%	50
HOG	79.42%	1,296

Table 1: Classification Performance Comparison

Analysis:

- **Highest Performance:** LBP (93.83%) - Best balance of accuracy and efficiency
- **Lowest Performance:** Raw Pixels (42.39%) - Insufficient for texture discrimination
- **Computational Efficiency:** LBP requires the fewest features (256) while achieving highest accuracy
- **Feature Effectiveness:** Texture-specific features (LBP, HOG) significantly outperform raw pixel intensities

The results demonstrate that texture-specific feature extraction methods are essential for effective texture classification, with LBP providing the optimal combination of accuracy and computational efficiency.