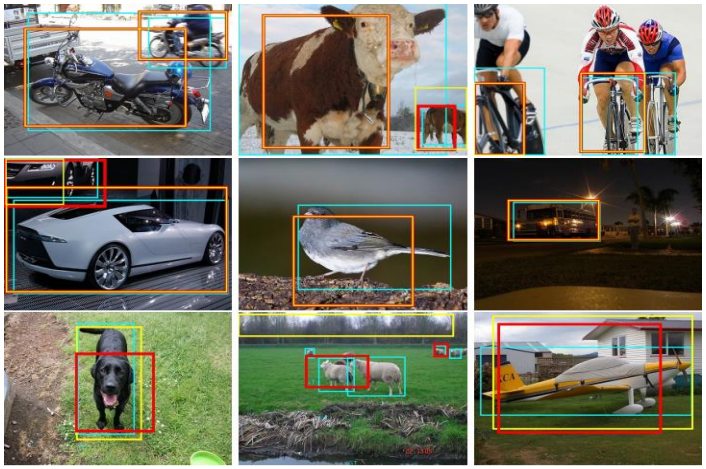
Designing Feature Descriptor for Image Classification



Presented By

Md. Mostafijur Rahman
MSSE0303
Institute of Information Technology
University of Dhaka

Supervised By

Dr. Mohammod Shoyaib
Associate Professor
Institute of Information Technology
University of Dhaka

Outline

- Introduction and Motivation
- Existing Works
- Problem Specification
- Proposed Solution
- Experimental Evaluation
- Conclusion and Future Work
- Q & A

Introduction

Image classification: Given a set of images, the objective is to classify those by the

- face,
- expression (e.g., happy, sad),
- gender (e.g., male, female),
- scene (e.g. coast, forest, classroom, kitchen),
- object categories (e.g. cars, leopards, laptop).

Scope

- Face verification
- Facial expression recognition
- Gender classification
- Scene classification
- Texture classification

- Object classification
- Aerial image classification
- Garments pattern classification
- Flower classification
- Leaf classification

Applications

- Biometric authentication
- Access control
- Surveillance system
- Market demand analysis
- Photosynthesis
- Medical imaging
- Entertaining tools

General Framework

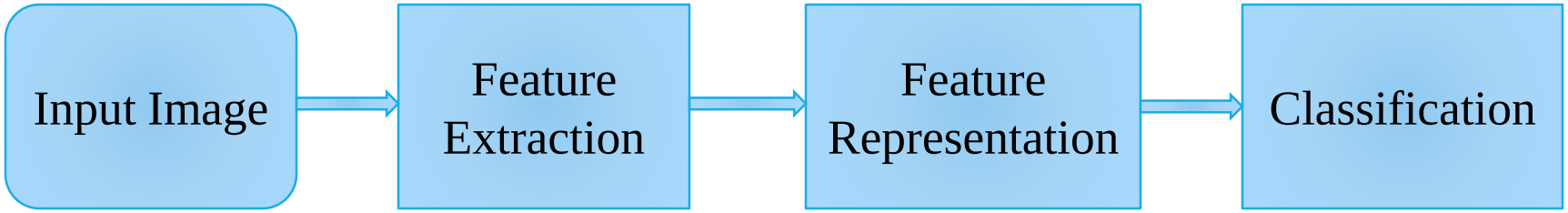


Fig. 1: General steps in classification systems using face image

General Framework

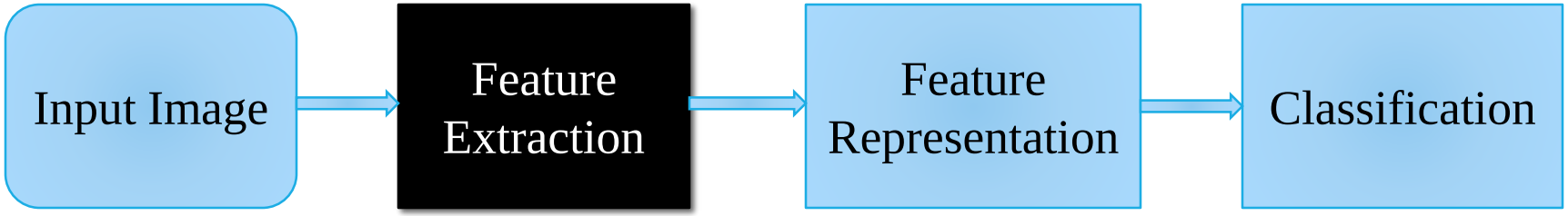


Fig. 1: General steps in classification systems using face image

Desired Properties

- Discriminating Ability
- Illumination Invariance
- Generalizability
- Stable Code

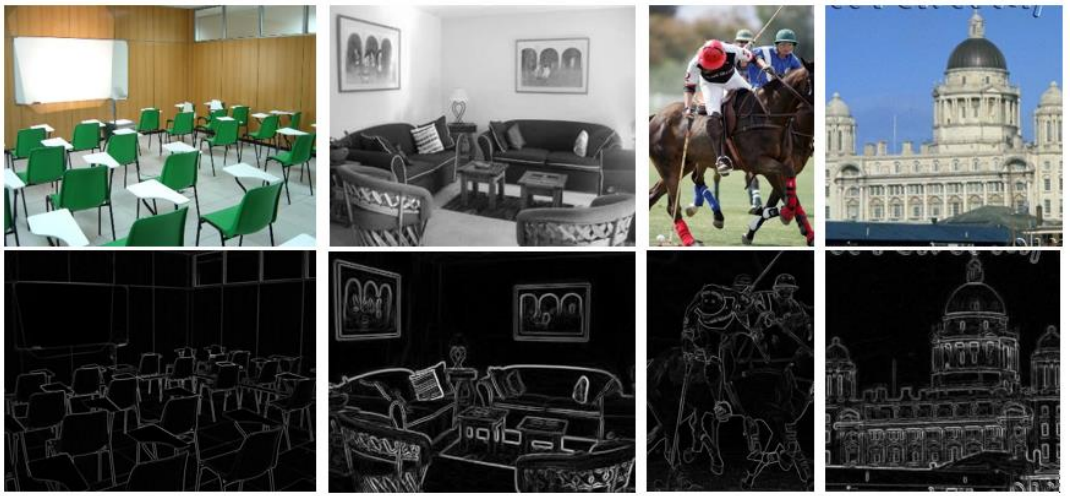


Fig. 2: Sample images with corresponding Sobel images from different categories of object

Existing Works

Gradient based

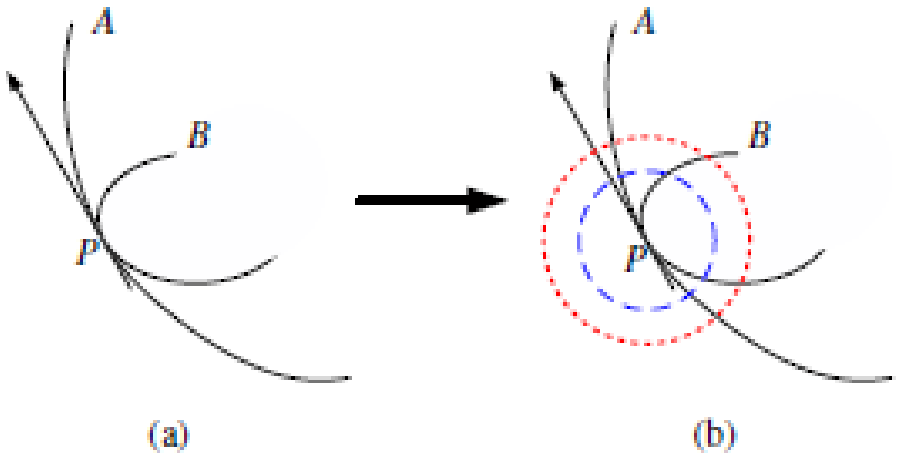
- Histogram of Oriented Gradient (HOG) [1]
- Scale Invariant Feature Transform (SIFT) [2]
- Gabor Filters [3]

LBP based

- Local Binary Pattern (LBP) [4, 5]
- Local Gradient Pattern (LGP) [6]
- Local Ternary Pattern (LTP) [7]
- Local Tetra Pattern (LTrP) [8]
- Local Direction Number Pattern (LDN) [9]
- Local Derivative Pattern (LDP) [10]

Problem Specification: gradient based techniques

- Two gradients having same direction may correspond to different local structures but gradient based methods fail to differentiate those.



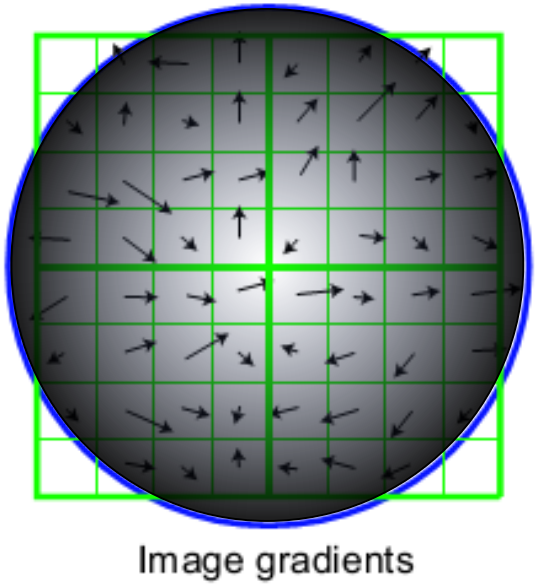


Fig. 3: Limitation of gradient based techniques [11]

Problem Specification: LBP/ CENTRIST

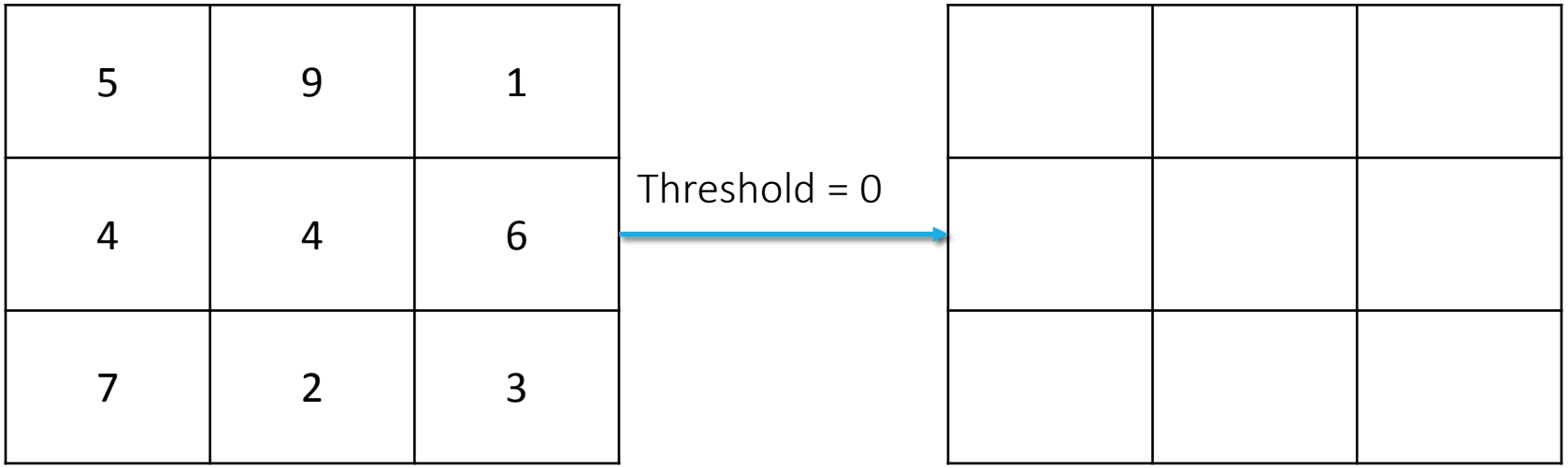


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

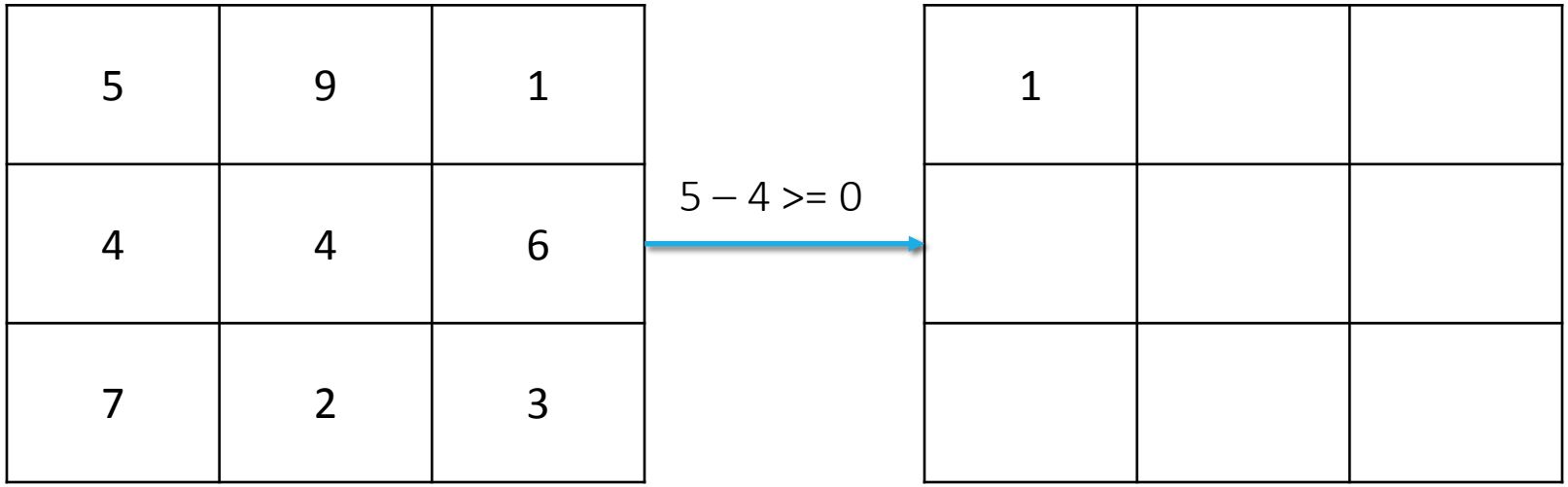


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

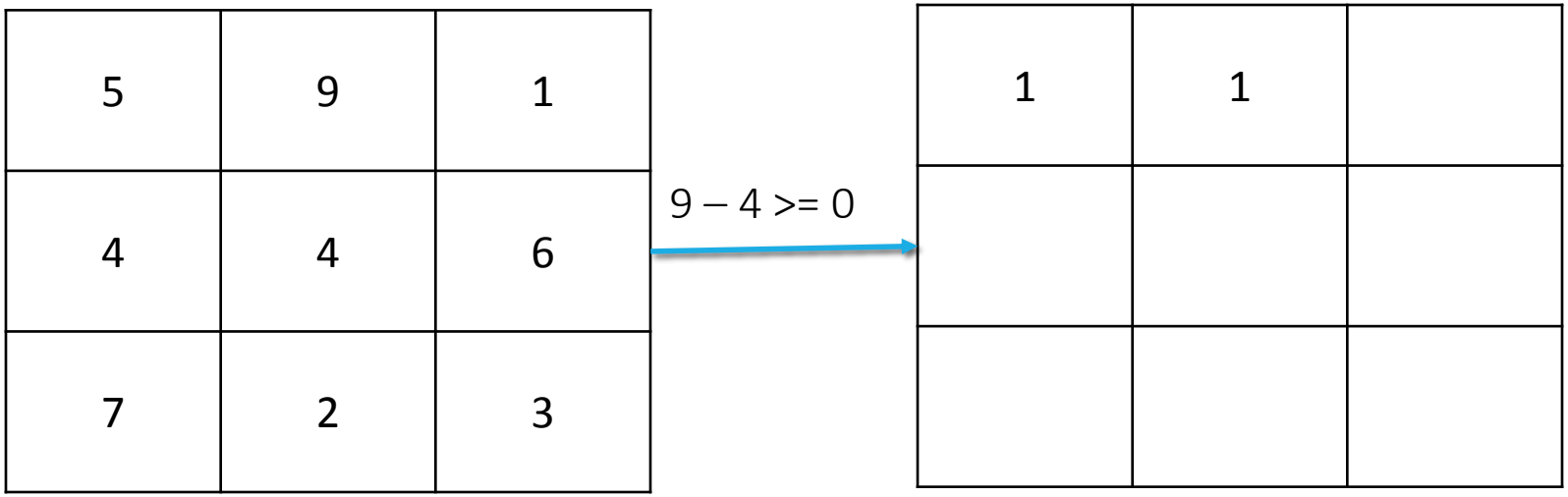


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

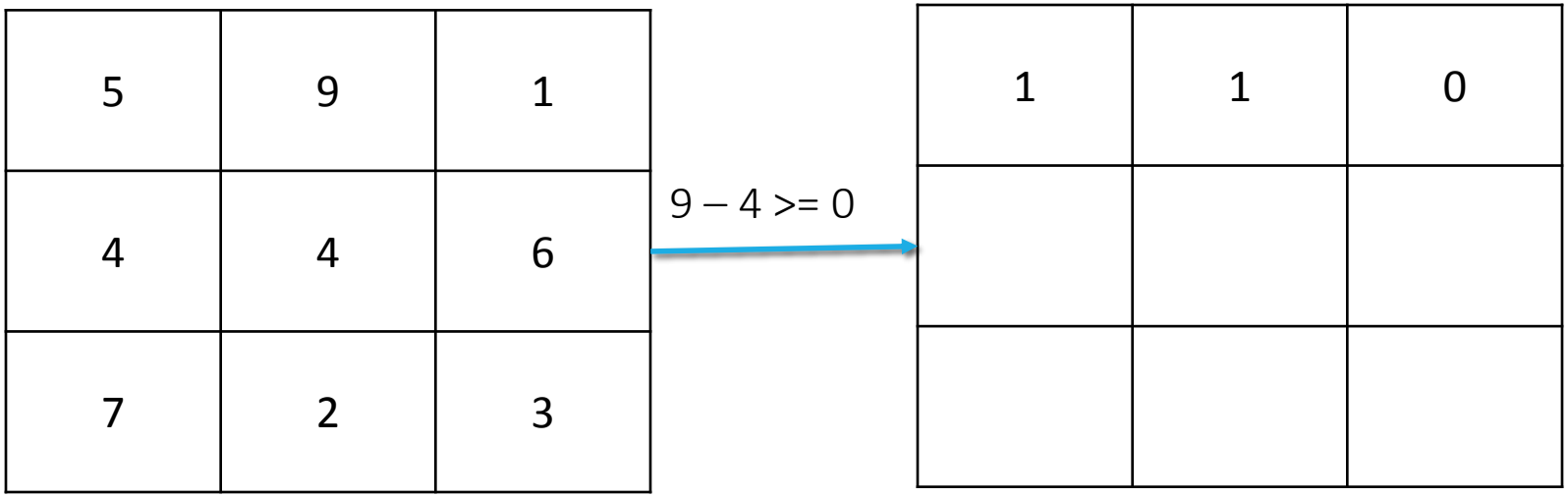


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

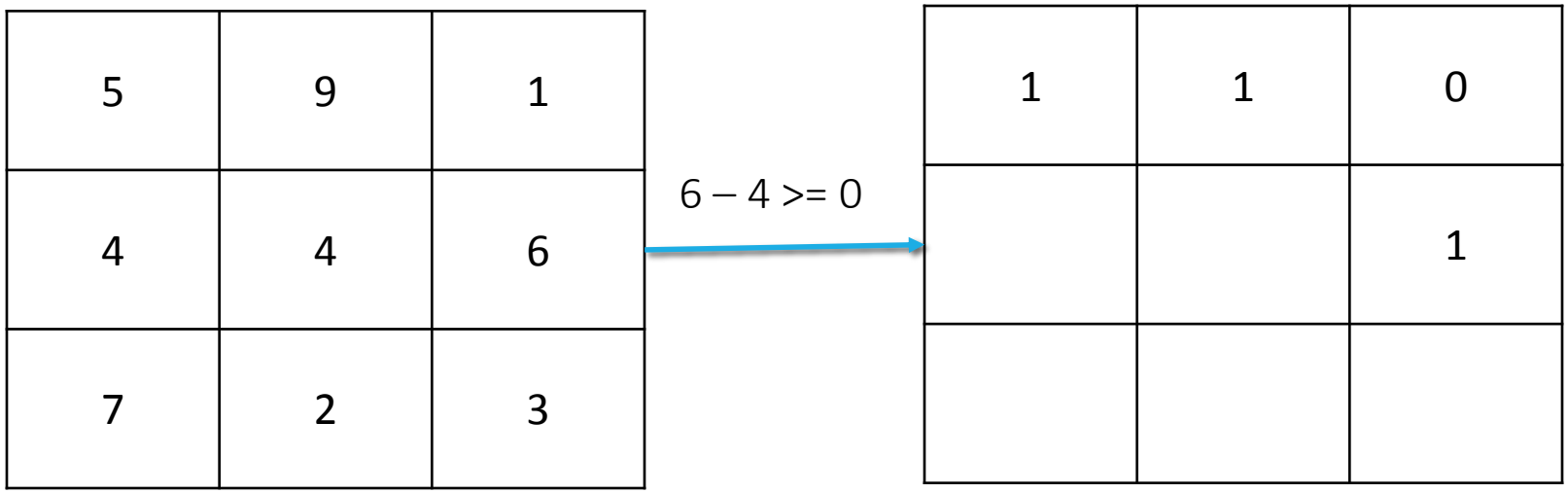


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

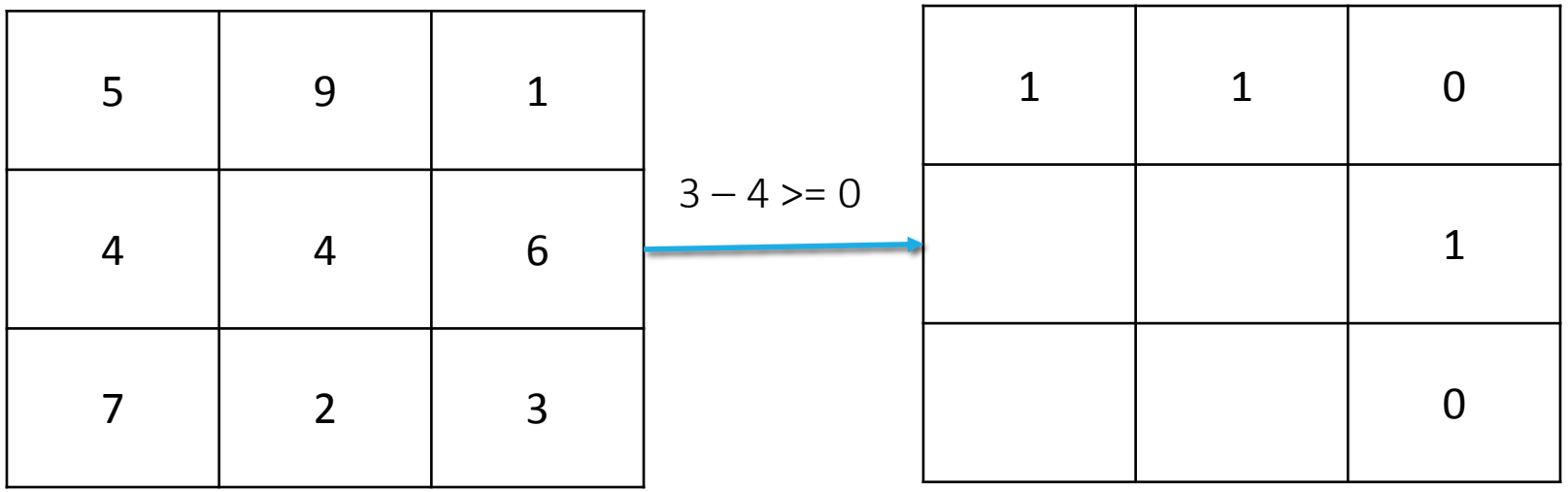


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

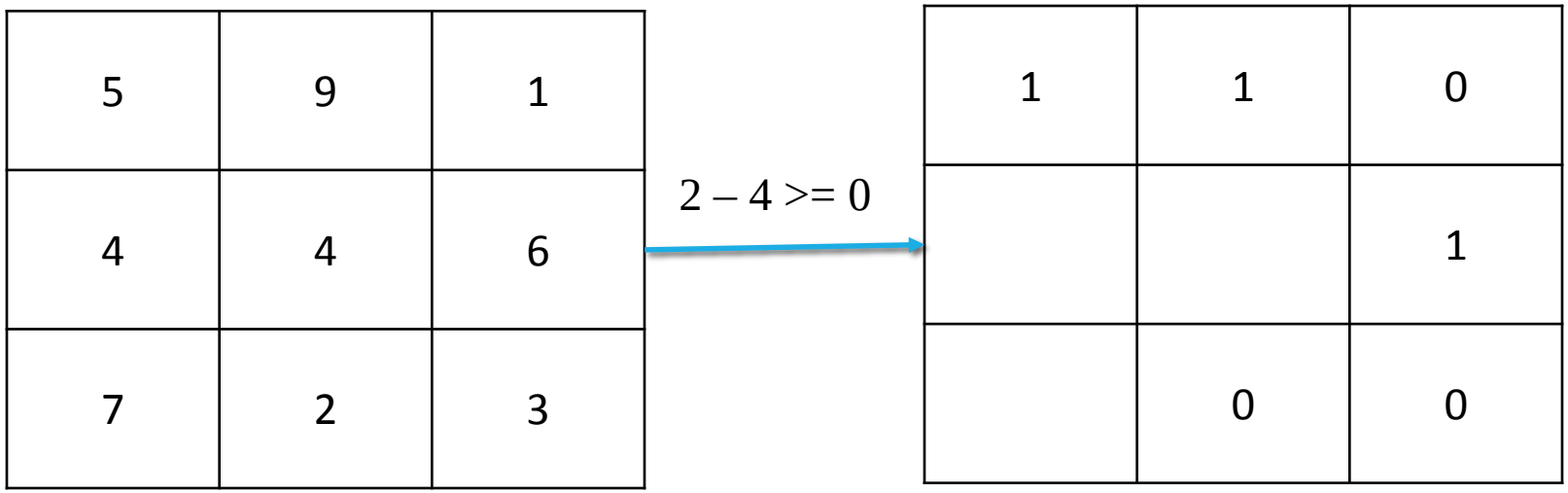


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

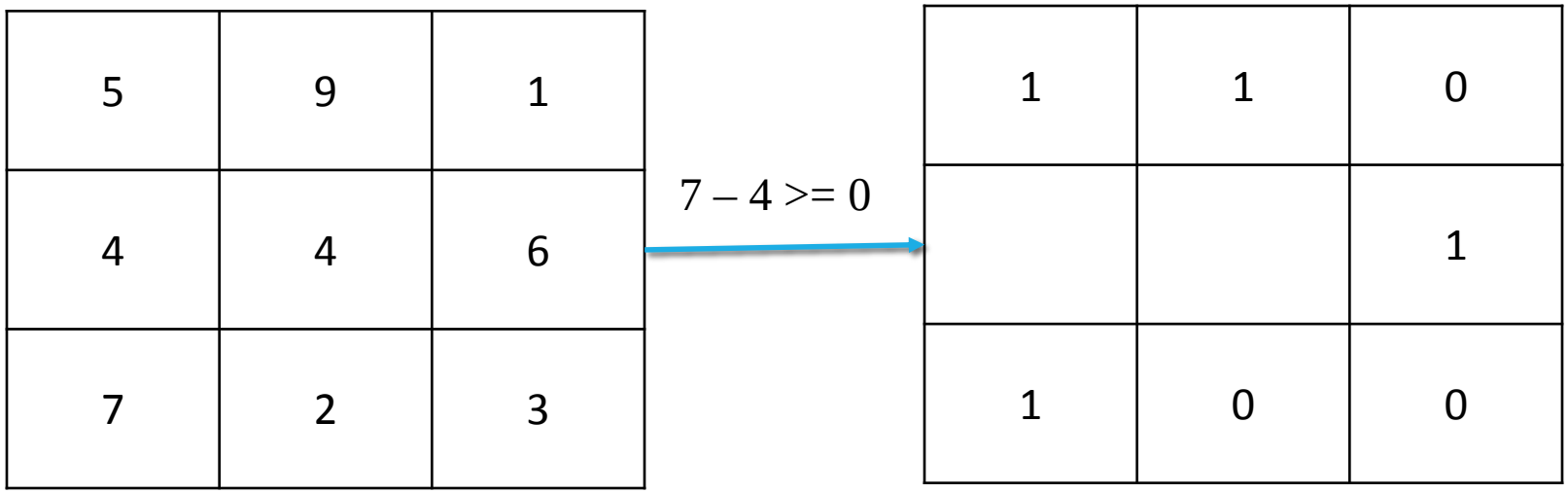


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

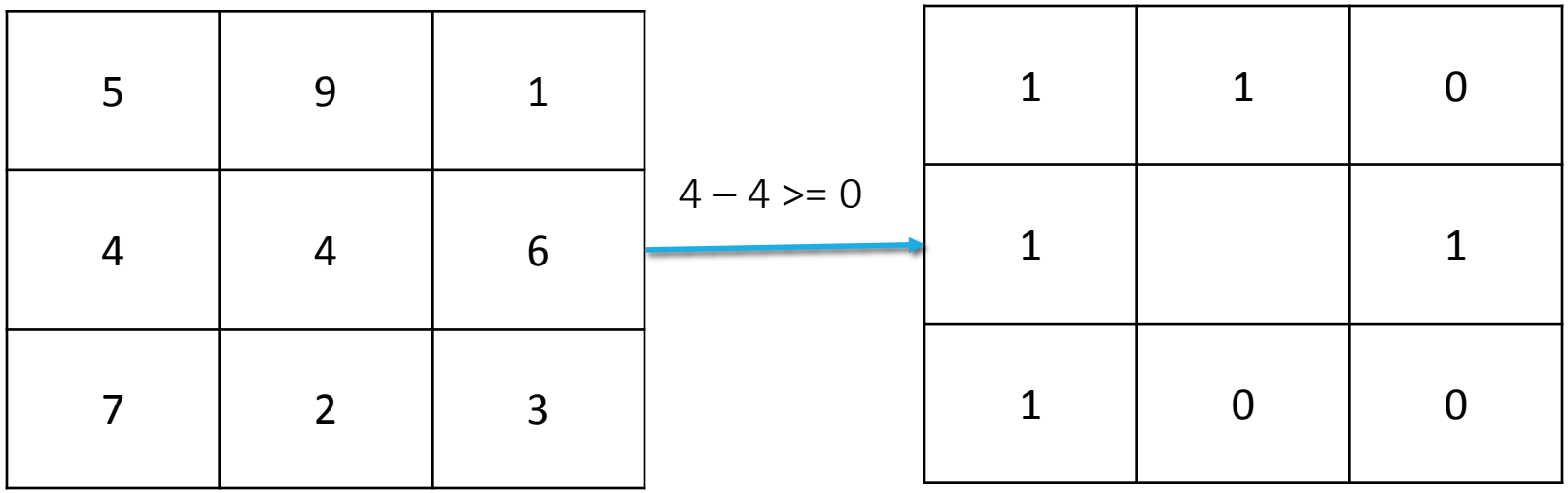


Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP/ CENTRIST

$(11010011)_2 = (211)_{10}$

5	9	1
4	211	6
7	2	3

Fig. 4: Basic LBP/CENTRIST

Problem Specification: LBP based techniques

- Lack of discriminating ability

171	174	175
173	170	172
174	175	171

(a)

$(11111111)_2 = (255)_{10}$

190	195	194
196	170	193
182	183	197

(b)

$(11111111)_2 = (255)_{10}$

Fig. 5: LBP produces same pattern for (a) small and (b) large differences

Problem Specification: LBP based techniques

- Sensitive to noisy intensity fluctuation

Fig. 6: Noise sensitivity of LBP and CENTRIST (a) Original texture, (b) texture change due to intensity fluctuation

Solution Direction

- Using other thresholds (i.e., 5) rather than zero

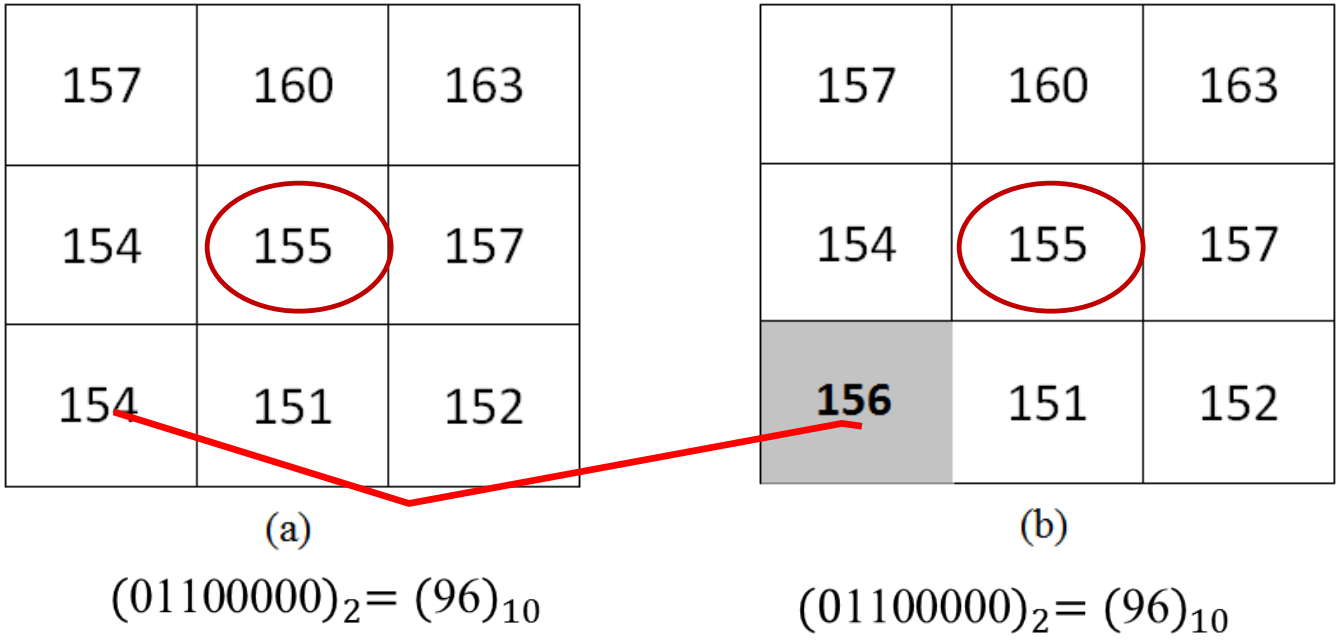


Fig. 7: (a) Original texture, (b) texture change due to intensity fluctuation

Noise Adaptive Binary Pattern (NABP)

NABP: Threshold Selection

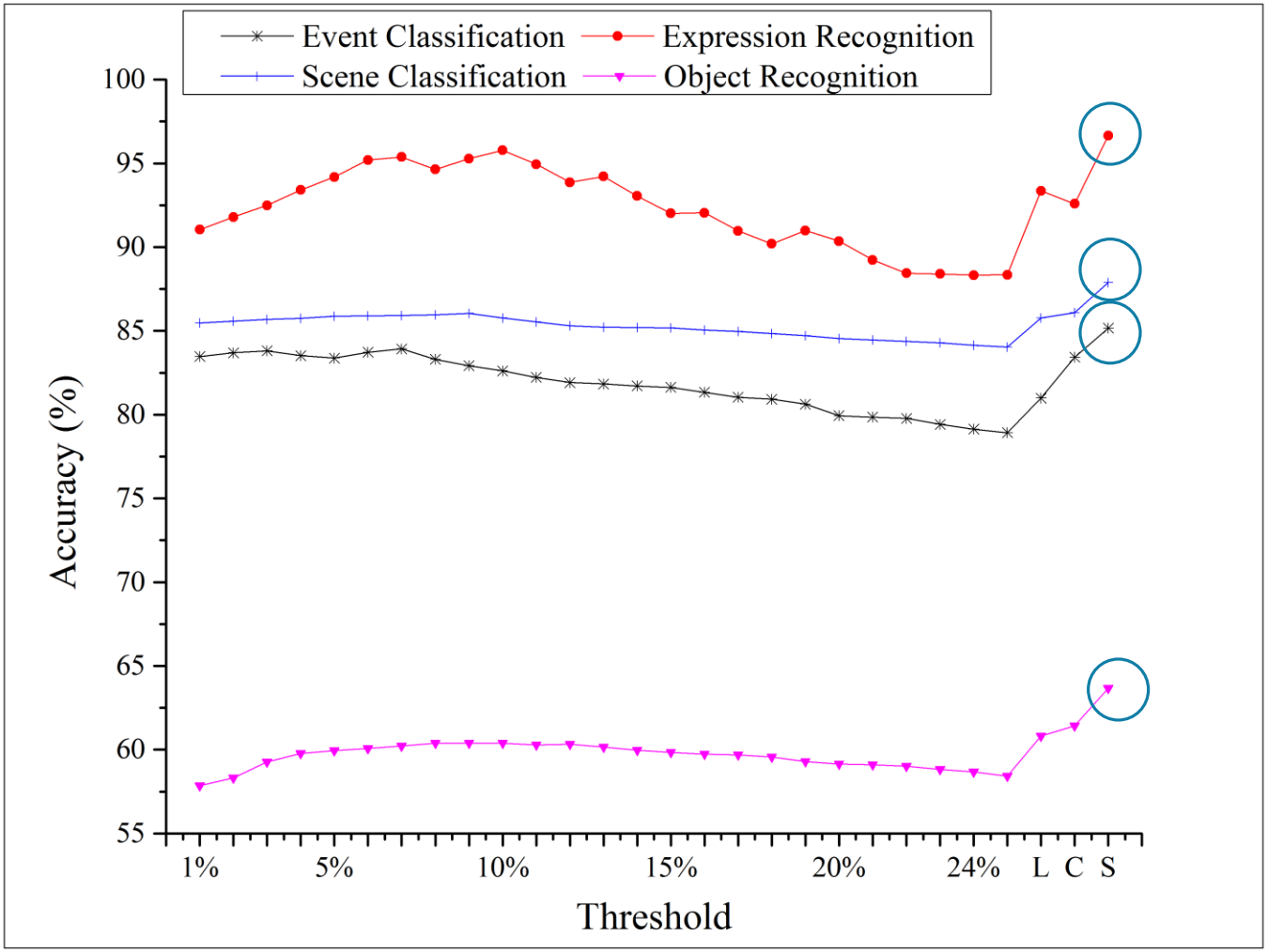


Fig. 8: Event, scene, object and expression recognition using different thresholds. L (threshold used in LGP), S (SQRT of center pixel) and C (cube root of center pixel) **(Adaptive threshold)**

NABP: Threshold Selection

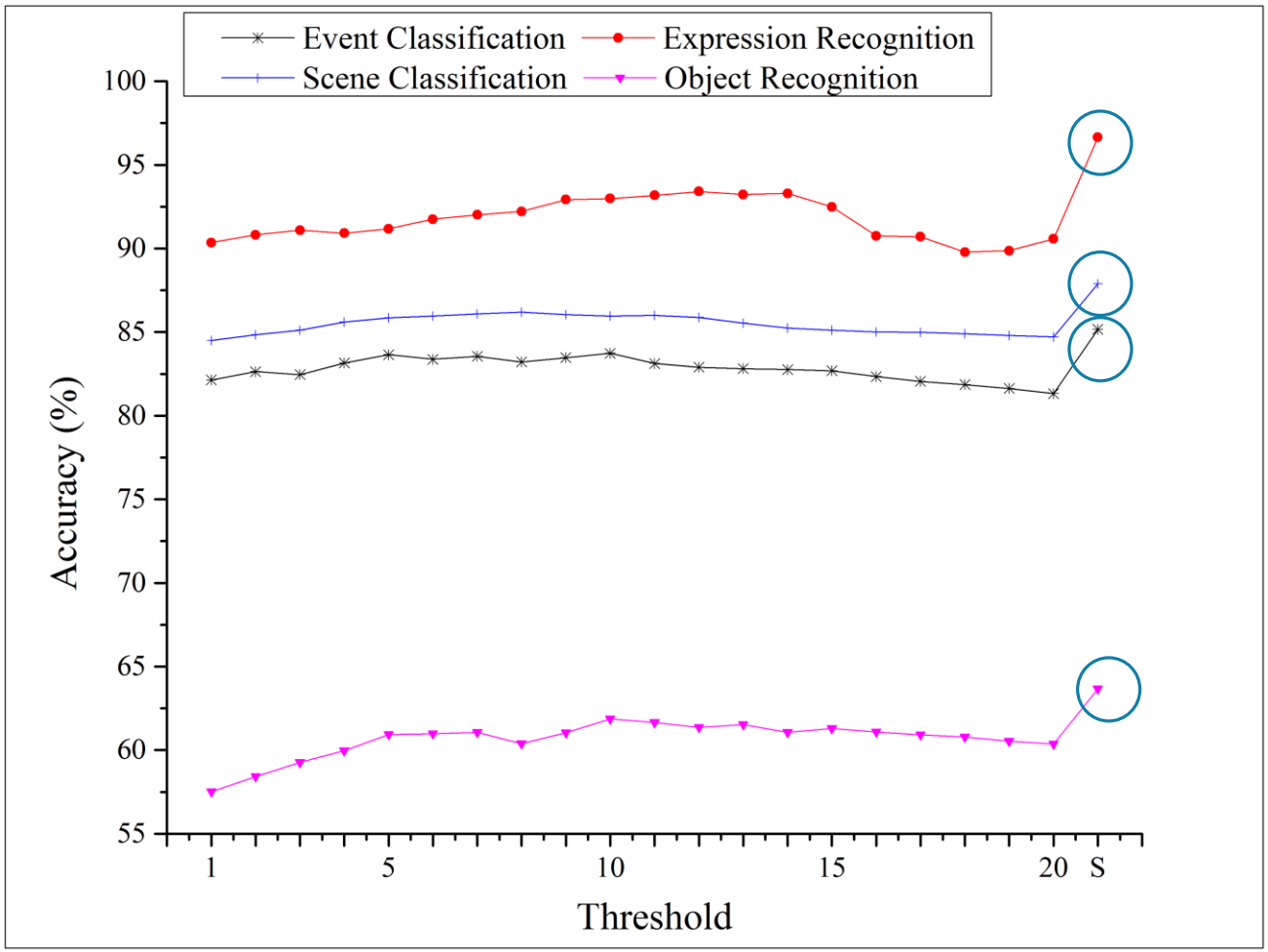


Fig. 9: Event, object, scene and expression recognition using different thresholds. Here, ***s*** is the **square root of center pixel** threshold. (Fixed threshold)

NABP

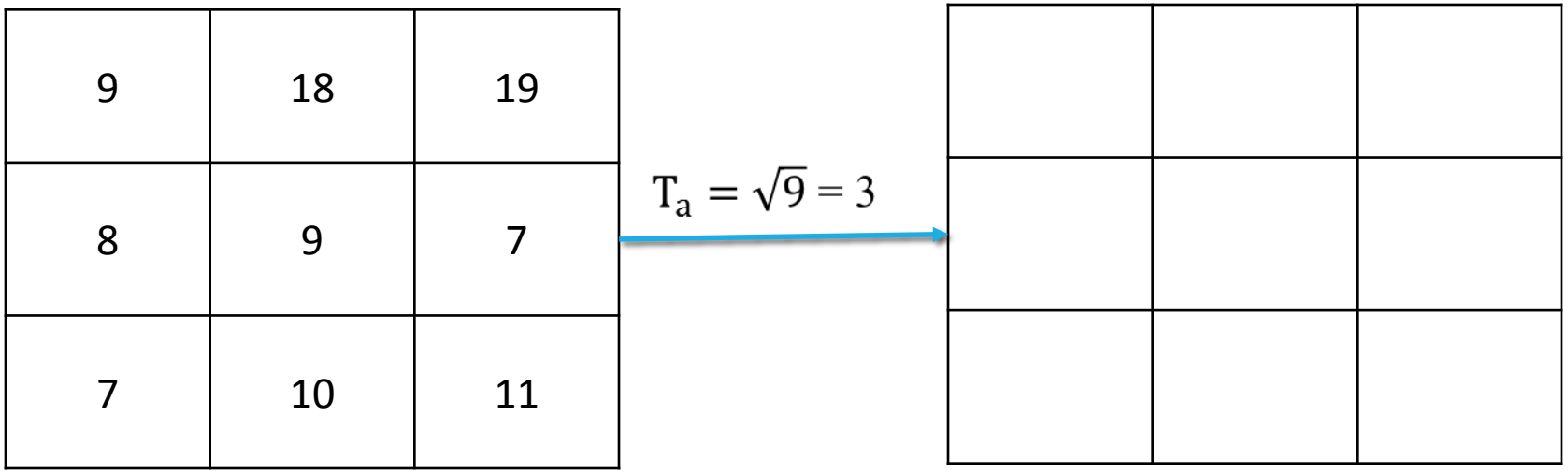


Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

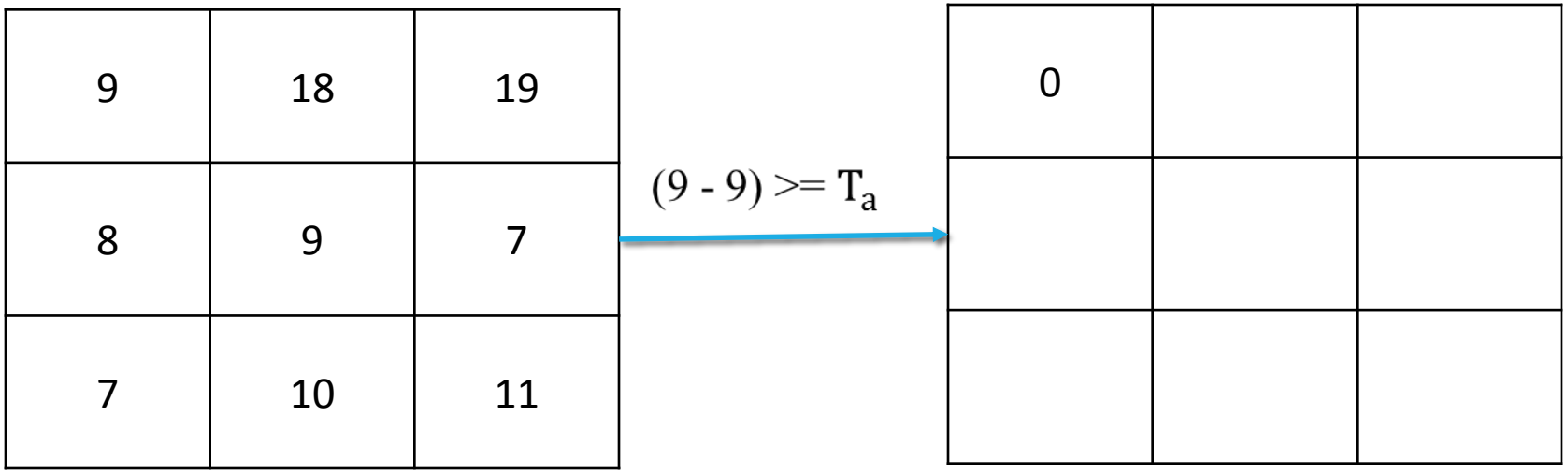


Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

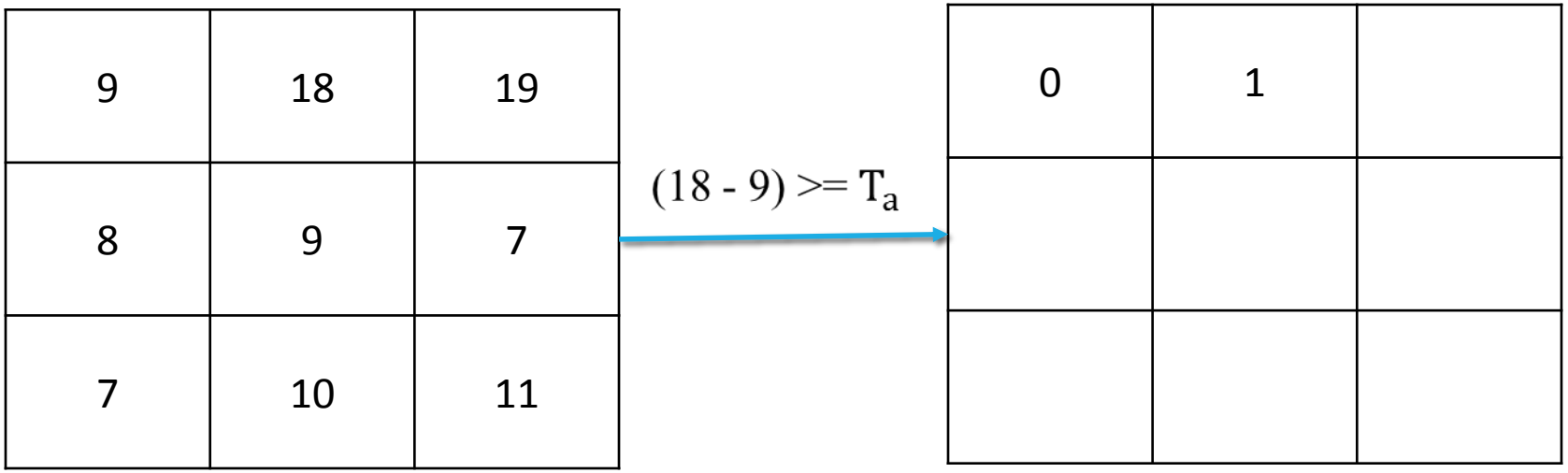


Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

9	18	19
8	9	7
7	10	11

$(19 - 9) >= T_a$

0	1	1

Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

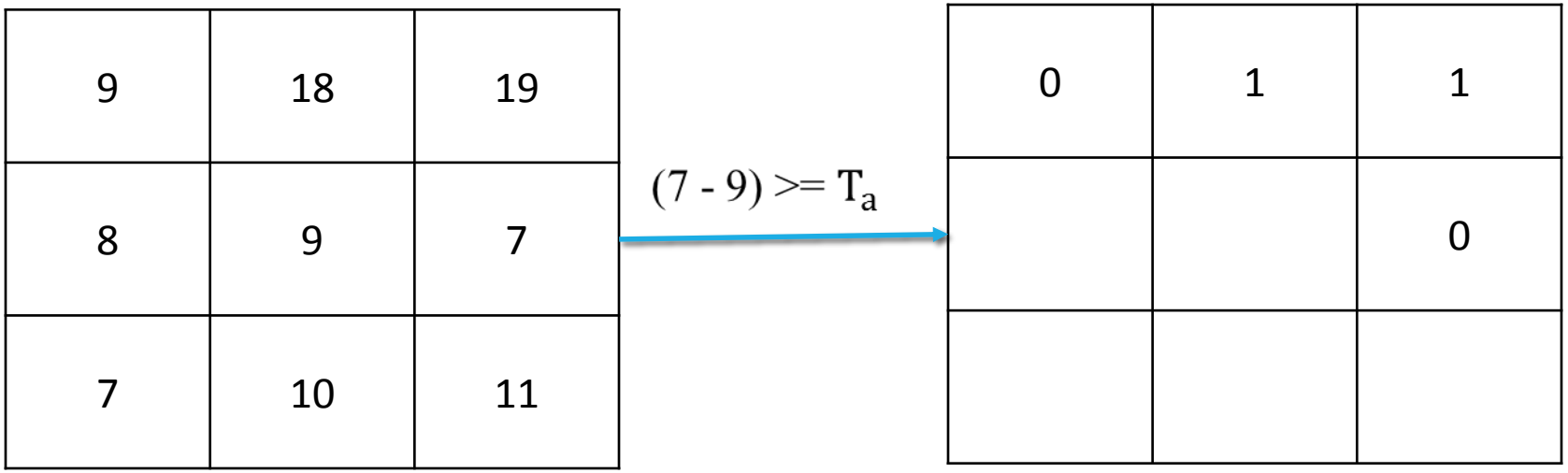


Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

9	18	19
8	9	7
7	10	11

$(11 - 9) >= T_a$

0	1	1
		0
		0

Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

9	18	19
8	9	7
7	10	11

$(10 - 9) >= T_a$

0	1	1
		0
	0	0

Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

9	18	19
8	9	7
7	10	11

$(7 - 9) >= T_a$

0	1	1
		0
0	0	0

Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

9	18	19
8	9	7
7	10	11

$(8 - 9) >= T_a$

0	1	1
0		0
0	0	0

Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

$(01100000)_2 = (96)_{10}$

5	9	1
4	96	6
7	2	3

Fig. 10: NABP coding process where $T_a = \sqrt{9} = 3$

NABP

Feature Representation

- Final feature is represented taking the histogram of coded image (quantized into only uniform bins)
- Here, uniform bins – at most two bit-wise transition (for 8 bits images total 59 bins)

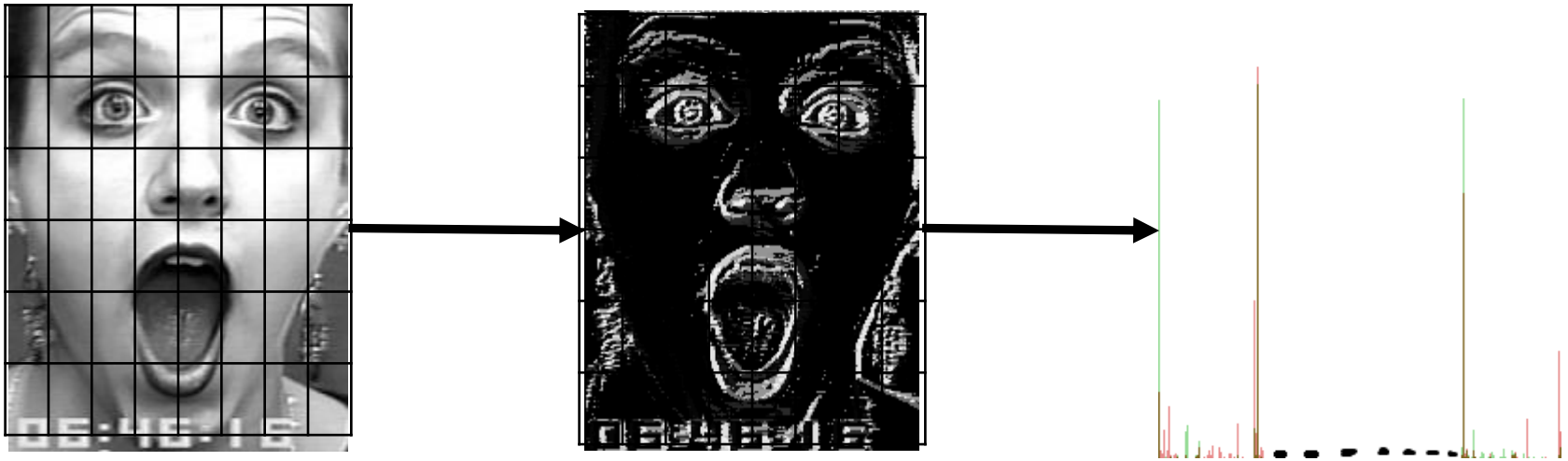


Fig. 11: NABP feature representation

Properties of NABP: Discriminating Ability

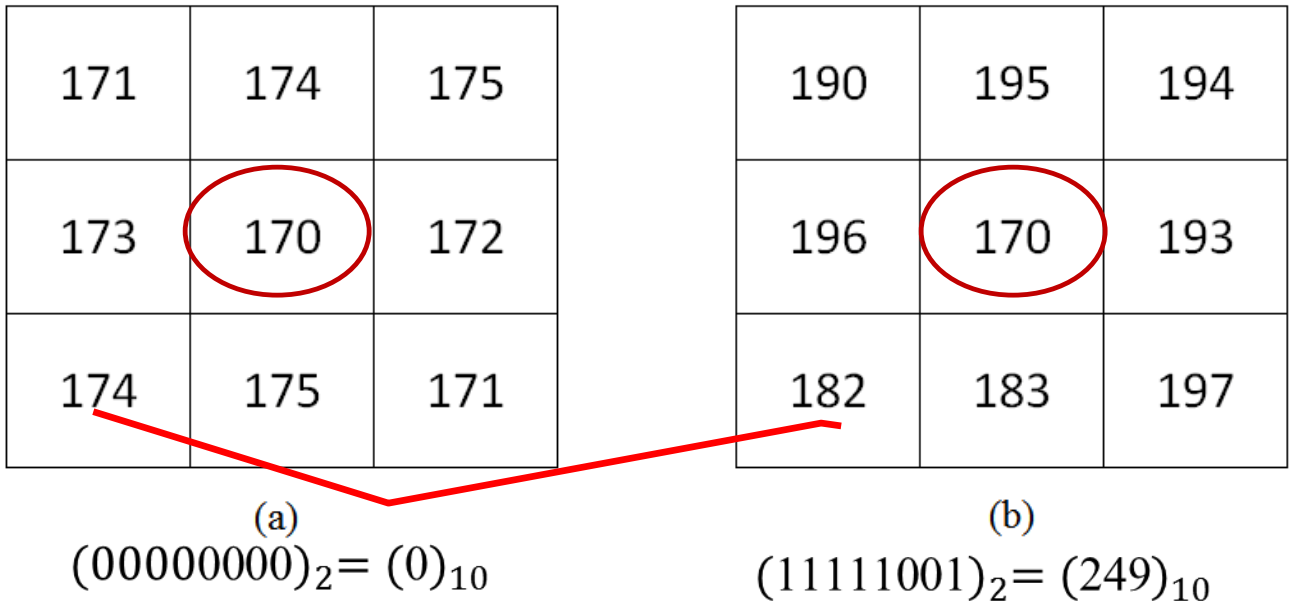


Fig. 12: NABP produces different patterns for (a) small and (b) large differences

Properties of NABP: Discriminating Ability

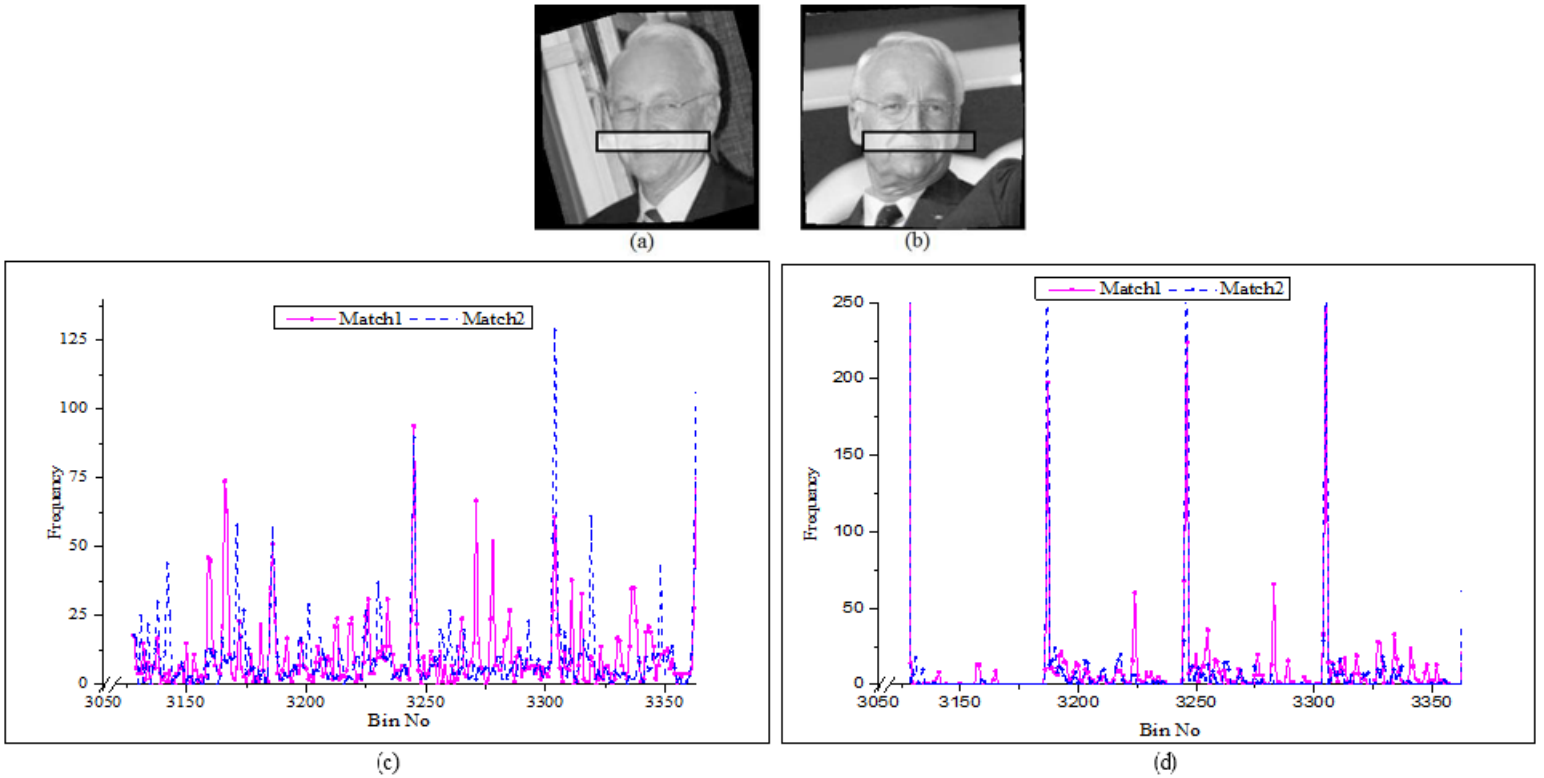


Fig. 13: Similarity of histograms for two images of a person (a) Match image 1, (b) Match image 2, (c)-(d) LBP and NABP histogram of (a) and (b) respectively

Properties of NABP: Discriminating Ability

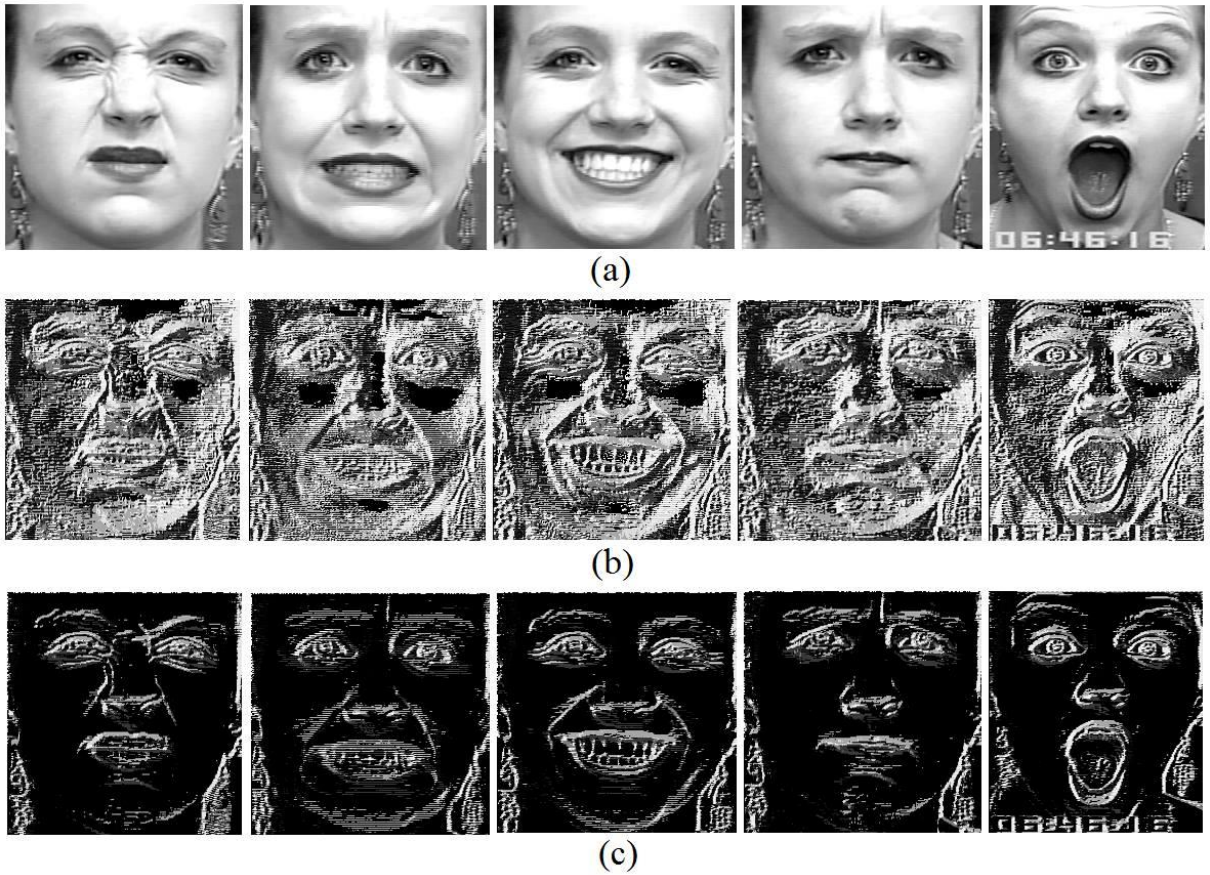


Fig. 16: Discriminative power of different methods in expression recognition, (a) original image, (b) LBP and (c) NABP

Properties of NABP: Noise Adaptive

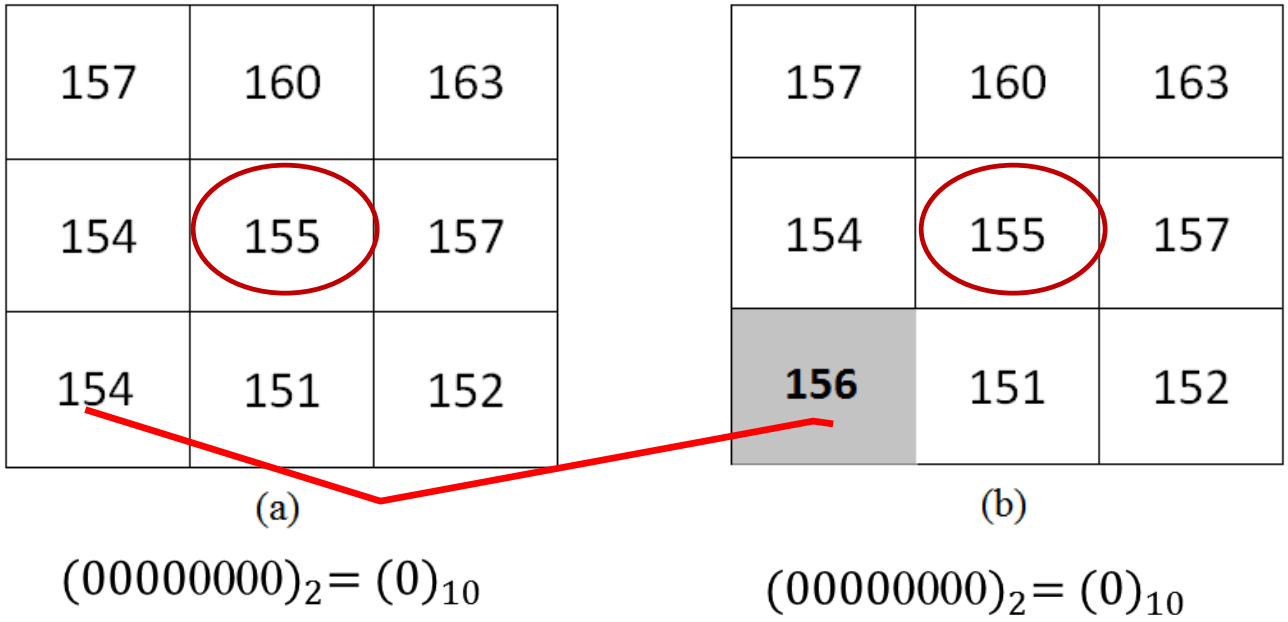


Fig. 18: Noise adaptiveness of NABP (a) original texture, (b) texture change for intensity fluctuation

Properties of NABP: Noise Adaptive

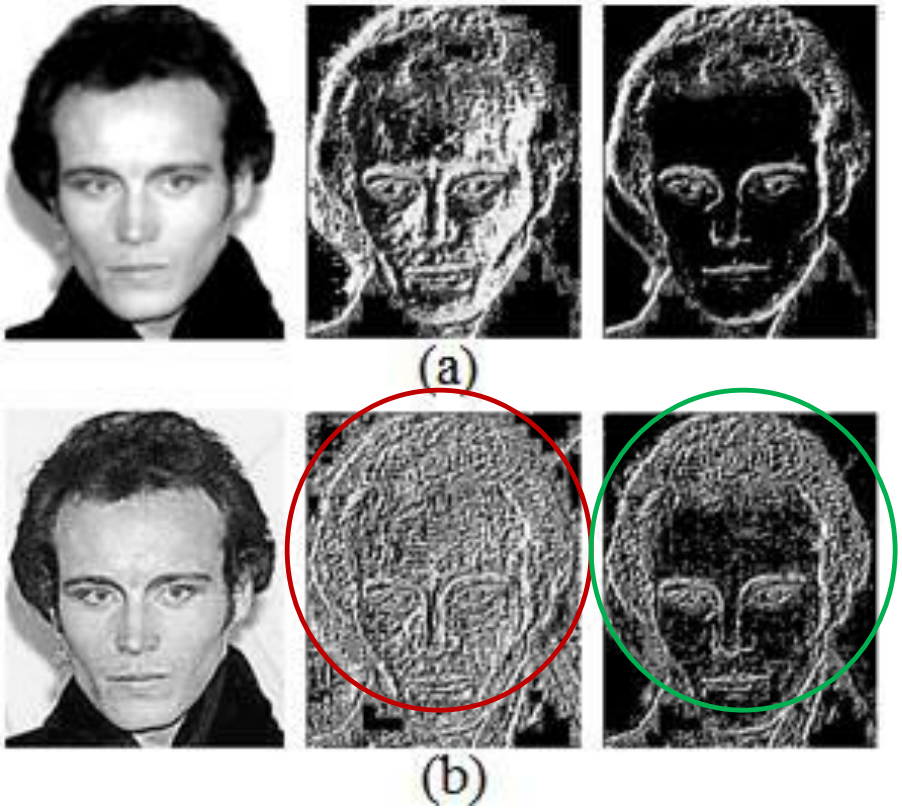


Fig. 19: Validating noise adaptiveness of NABP (a) original texture, (b) noisy texture (left column original image, middle column LBP coded image and right column NABP coded image)

Limitation of NABP

- NABP cannot discriminate between a positive and negative changes with the center pixel.
- Only experimental justification of the threshold selection

Discriminating Ternary Census Transform Histogram (DTCTH)

DTCTH: Threshold Selection

- Small threshold in small intensity and large threshold in high intensity
- Capture the salient textures in both small and high intensity region
- By analyzing different types of noises
  - Shot noise is unavoidable noise
  - It is dominant noise in low light condition [13]
  - The expected magnitude of shot noise is ***square root*** of the intensity of photon [13]

DTCTH

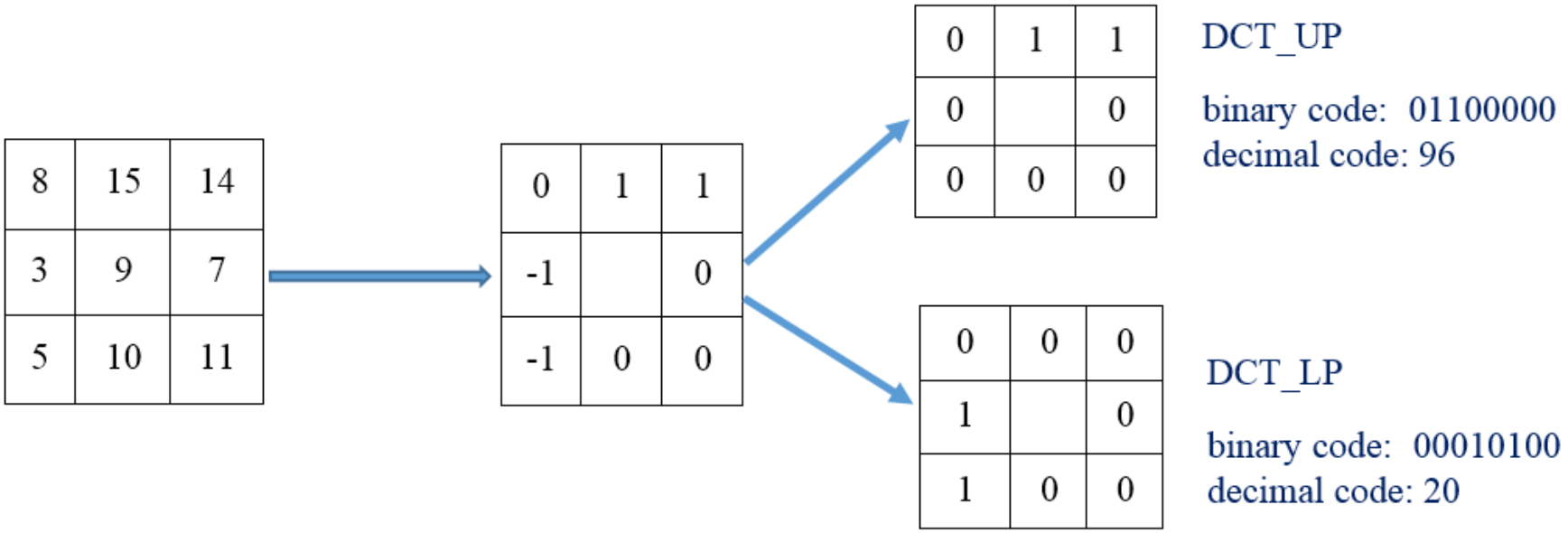


Fig. 20: DCT code generation

DTCTH

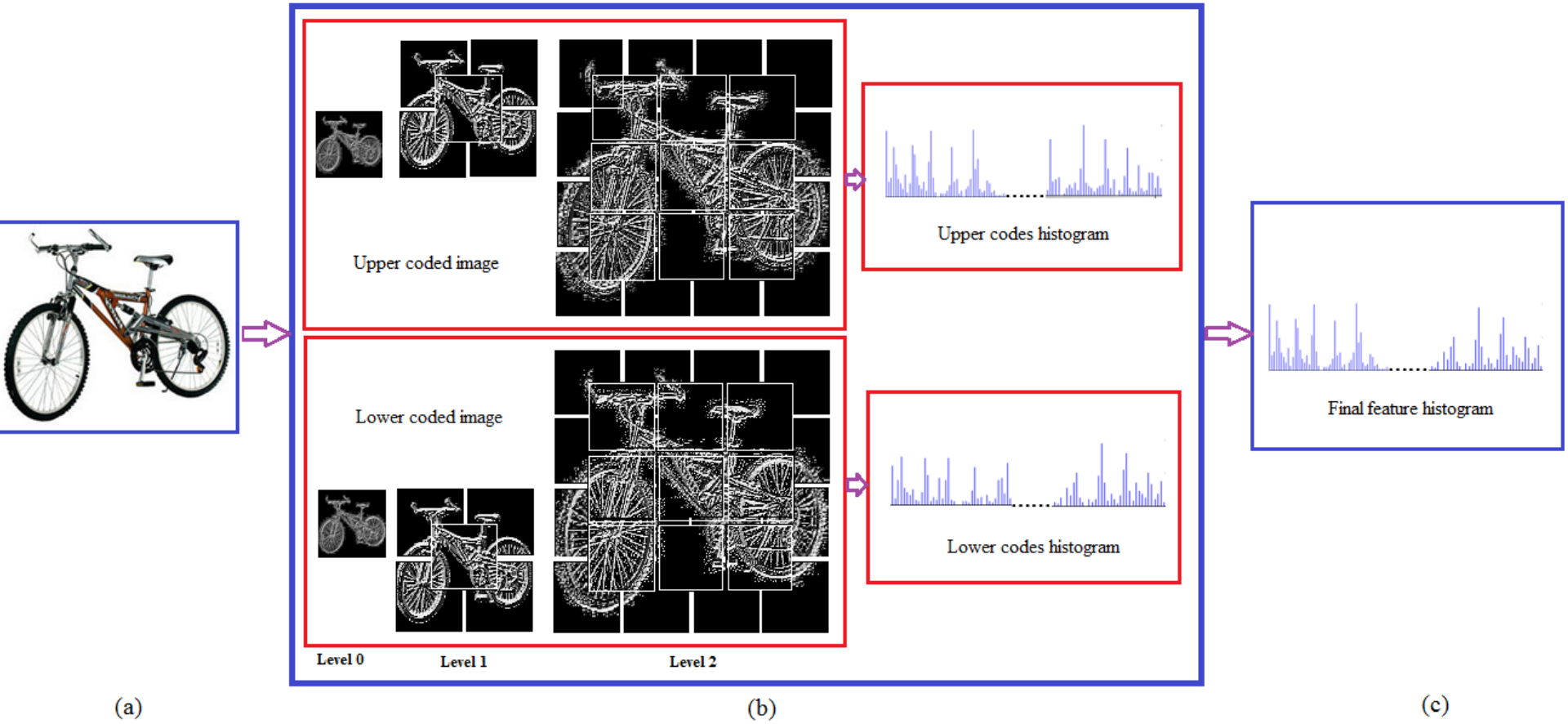


Fig. 21: Overall DTCTH feature generation, (a) input image, (b) spatial pyramid representation, (c) Final feature histogram

Properties of DTCTH

- Suppress background information and capture foreground information
- Produce stable code

Properties of DTCTH

- Discriminating ability for certain intensity changes in positive and negative directions

- Three groups of codes such as uncertain state (i.e., 0 for 72 and 69),
- intensity changes in positive direction (i.e., 1 for 77, 78 and 76) and
- negative direction (i.e., -1 for 64, 62 and 62) in certain regions by considering 70 as the center pixel.

77	78	76
72	70	69
64	62	62

Fig. 24: Example of discriminating code generation

Limitation of DTCTH

- DTCTH only considers sign information
- Does not consider color information

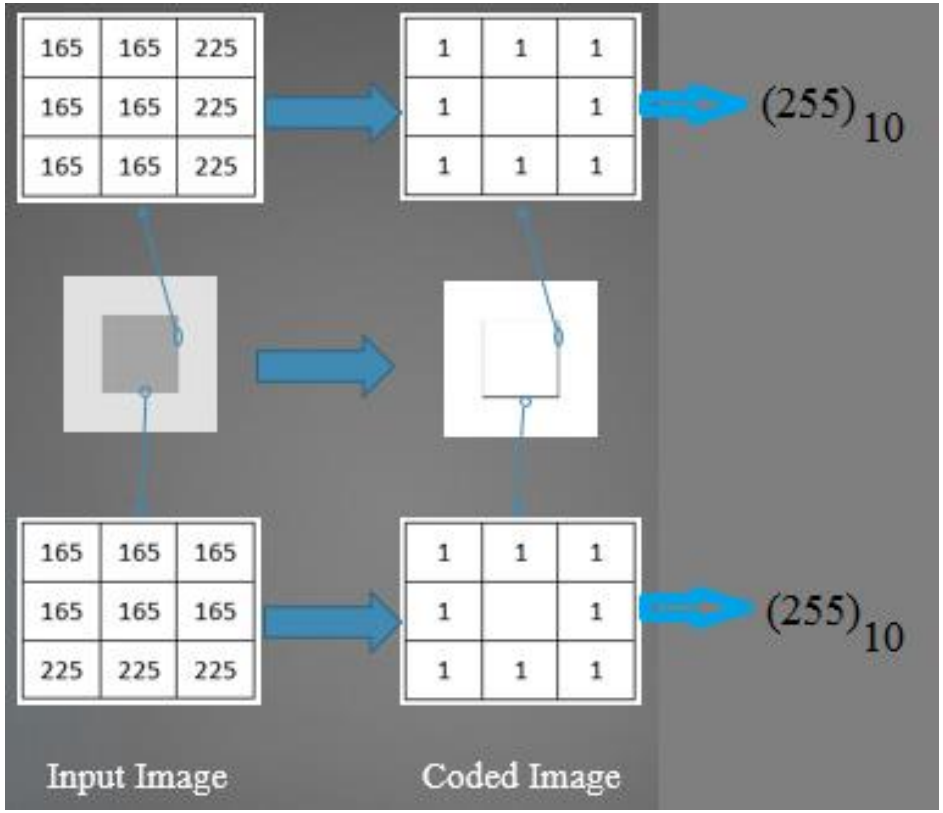


Fig. 25: Lack of discriminating ability of CENTRIST and LBP based descriptors

Multi-channel Complementary Census Transform (MCCT)

Analysis on Complementary Information

$$C_\mu = \frac{1}{K}\sum_{i=1}^{K}\frac{\sum_{j=1}^{N} d_{ij}}{D_i},\ d_{ij} = \begin{cases} 1, & I_{1j} \neq I_{2j} \\ 0, & otherwise \end{cases}$$

C_μ between Sign (S) and Magnitude (M) in Different Channels

	O1		O2		O3		Sobel_R	
	S	M	S	M	S	M	S	M
S	0.000	0.692	0.000	0.851	0.000	0.965	0.000	0.974
M	0.692	0.000	0.851	0.000	0.965	0.000	0.974	0.000

Analysis on Complementary Information

Average C_μ for Sign (S) Information of Different Opponent Channels

	O1	O2	O3	Sobel_R
O1	0.0000	0.8917	0.9260	0.9328
O2	0.8917	0.0000	0.8669	0.9190
O3	0.9260	0.8669	0.0000	0.9136
Sobel_R	0.9328	0.9190	0.9136	0.0000

Analysis on Complementary Information

Average C_μ for Magnitude (M) Information of Different Opponent Channels

	O1	O2	O3	Sobel_R
O1	0.0000	0.5959	0.6033	0.9031
O2	0.5959	0.0000	0.2963	0.9390
O3	0.6033	0.2963	0.0000	0.9434
Sobel_R	0.9031	0.9390	0.9434	0.0000

Analysis on Complementary Information

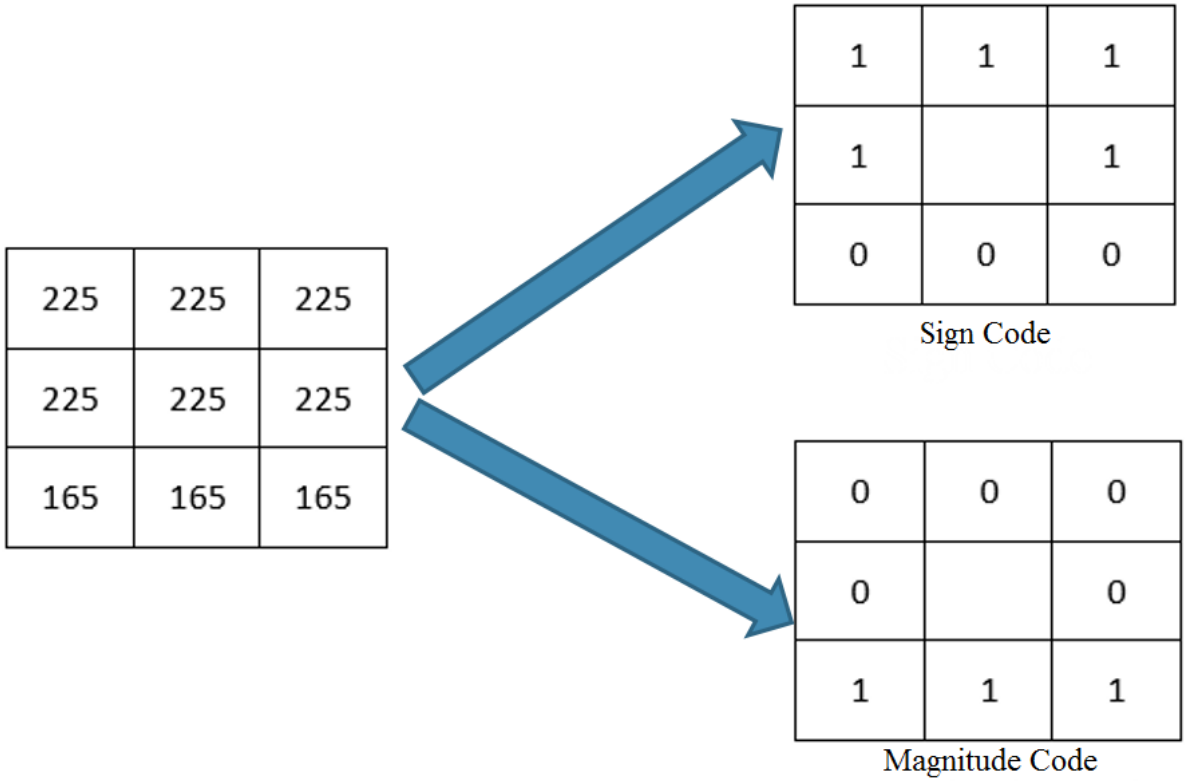


Fig. 26: An example of complementary property of sign and magnitude

MCCT

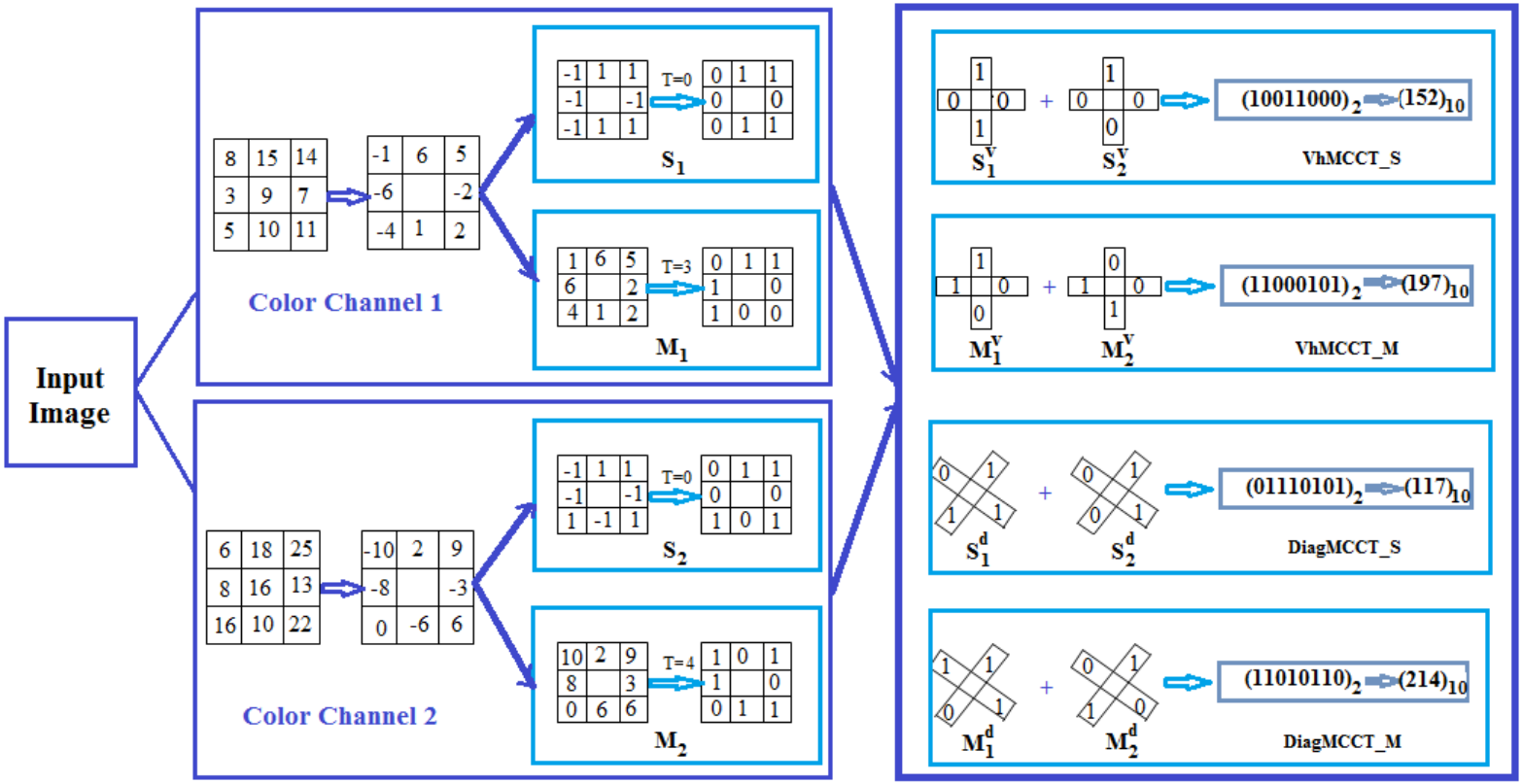


Fig. 27: MCCT code generation process

Properties of MCCT: Discriminating Ability

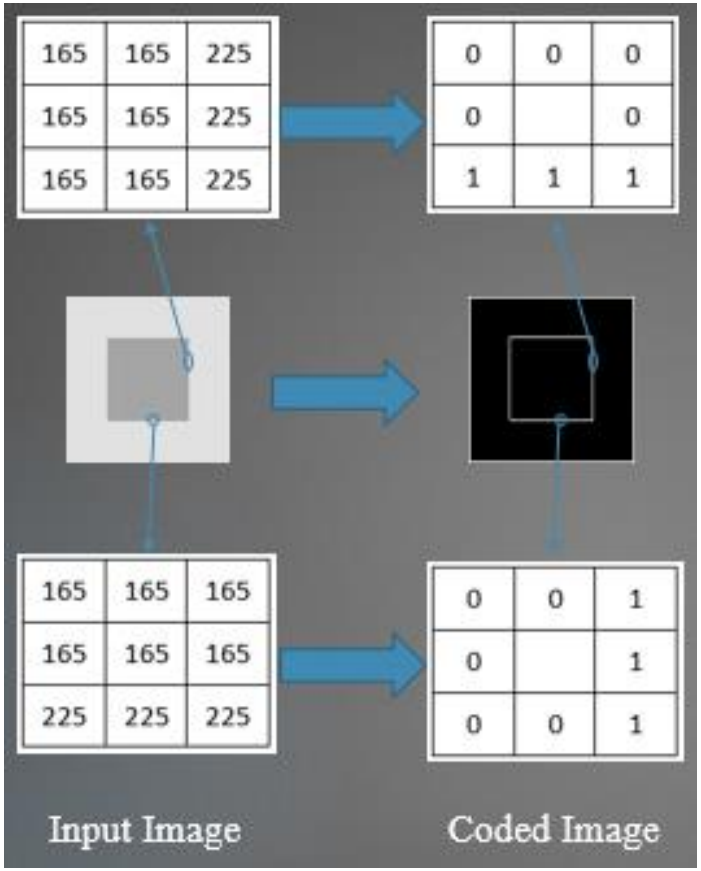


Fig. 28: Discriminating ability of magnitude in case of failure of sign information

Limitation of MCCT

- Capture uncertain regions (probable to be affected by intensity fluctuation) in a single code

Local Quaternary Census Transform (LQCT)

LQCT

$$LQCT_{n,r}(x_c, y_c) = \sum_{l=0}^{n-1} q(g_l) \times 4^l$$

$$q(g) = \begin{cases} 0, & if\ g \geq T \\ 1, & if\ g \leq -T \\ 2, & if\ 0 < g < T \\ 3, & if\ -T < g < 0 \end{cases}$$

Here, $g_l = |p_l - p_c|$ and $T = \sqrt{(p_c)}$

LQCT

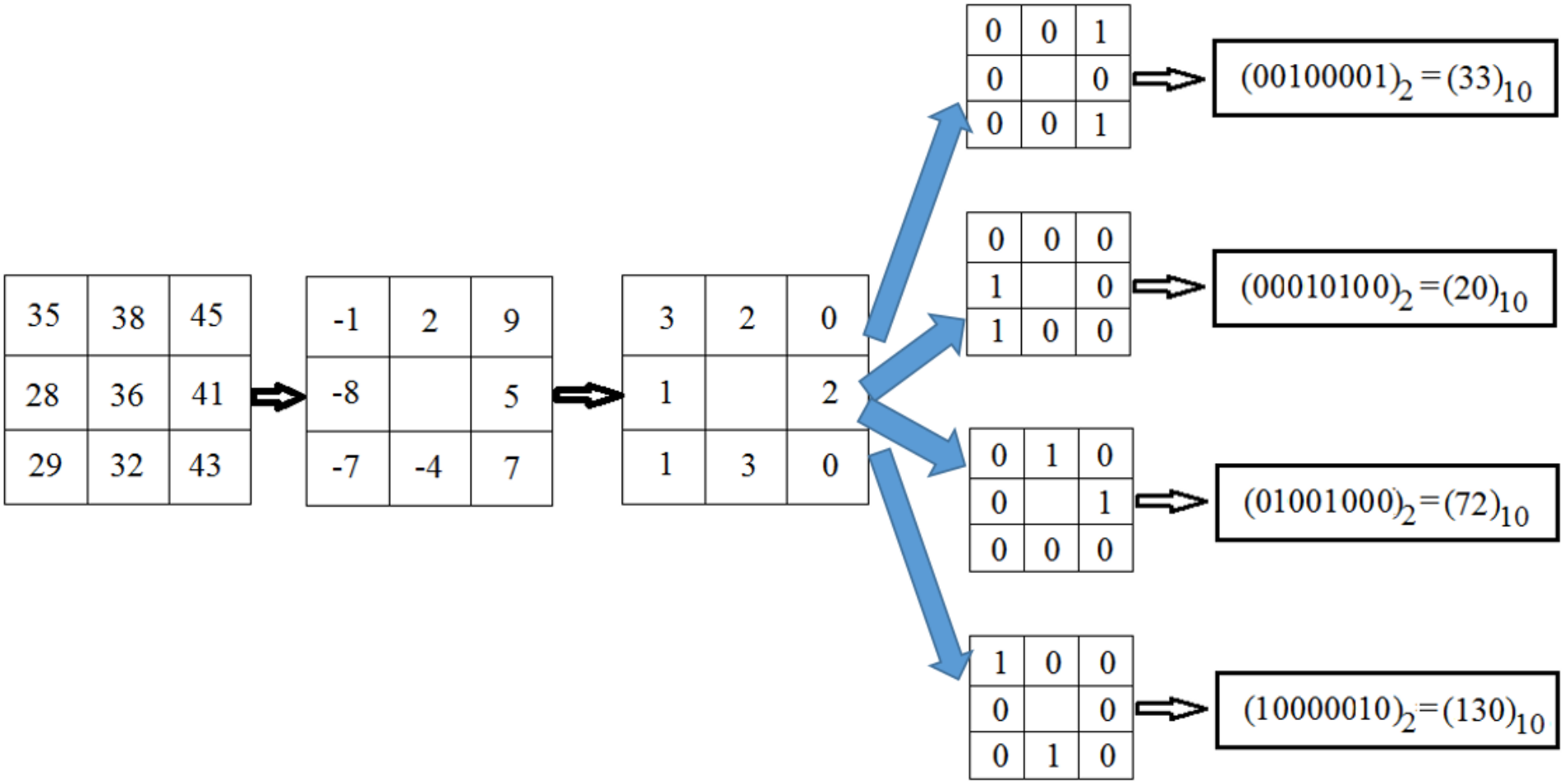


Fig. 29: LQCT code generation process

Multi-channel LQCT (mLQCT)

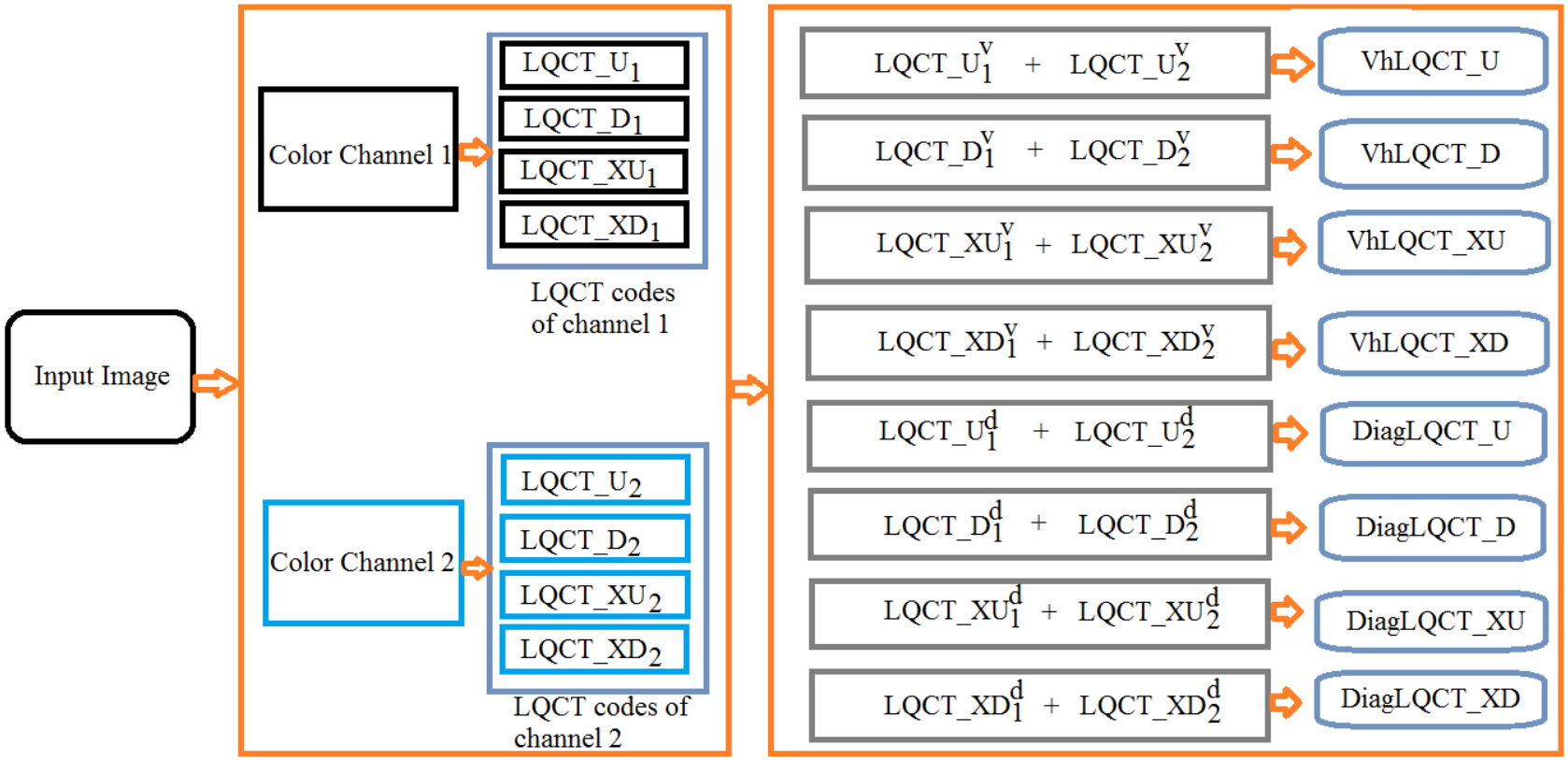


Fig. 30: mLQCT code generation, (a) LQCT codes from two different channels, (b) code combining process in these channels

Experimental Analysis

Implementation Details

- CENTRIST and mCENTRIST frameworks
- Remove the 0^{th} and 255^{th} bins from the histogram
- Take the square root of features
- Adaboost classifier for face recognition
- SVM classifier with Linear and Histogram Intersection Kernels for other applications

Experimental Evaluation: Dataset

Applications	Datasets
Face recognition	LFW View 2
Facial expression classification	CK, CK+
Gender classification	LFW
Flower classification	Oxford Flower 102
Scene classification	MIT Indoor 67, OT scene, Scene 15, RGB-NIR scene, SUN 397
Object classification	Caltech 101, Caltech 256
Event classification	UIUC Sports Event
Aerial image classification	Land-Use 21
Leaf classification	Swedish Leaf
Garments texture classification	Fashion, Clothing Attribute
Texture classification	FMD, K'th Tips
Skin Disease classification	3-Skin Disease

Experimental Evaluation: Dataset

Applications	Face Recognition	Facial Expression Classification		Gender Classification	Skin Disease Classification
Databases	LFW (View2)	Cohn Kanade (CK)	CK+	Color LFW	3-Skin Disease
Classes	Multi-class	6/7	7	2	3
Total Samples	12,000 (3,000 match pair 3,000 non-match pair)	960/1,280	981	13,230	446
Train Samples/ class	9 out of 10 folds	9 out of 10 folds		4 out of 5 folds	90
Test Samples/ class	1 out of 10 folds	1 out of 10 folds		1 out of 5 folds	Remaining

Experimental Evaluation: Dataset

Applications	Flower Classification		Event Classification	Leaf Classification	Aerial Image Classification
Databases	Oxford Flower 102	Flower 17	Sports Event 8	Swedish Leaf	Land-Use 21
Classes	102	17	8	15	21
Total Samples	8,189	1,360	1,586	1,125	2,100
Train Samples/ class	30	40	70	25	80
Test Samples/ class	Remaining	Remaining	60	Remaining	20

Experimental Evaluation: Dataset

Applications	Object Classification		Scene Classification				
Databases	Caltech-256	Caltech-101	OT Scene	Scene 15	Indoor 67	9 RGB-NIR Scene	SUN 397
Classes	257	102	8	15	67	9	397
Total Samples	30,608	9,145	2,688	4,485	5,620	477	1,08,574
Train Samples/class	60	30	100	100	80	42	50
Test Samples/class	Remaining	Remaining	Remaining	Remaining	20	Remaining	50

Comparison of Computational Overhead (Avg. Computation Time)

Techniques	Average Computation Time (sec)
ScSPM	43.95049
LLC	**45.29133**
CENTRIST	0.089131
mCENTRIST	0.880342
LBP	0.041313
LTP + Pyramid	0.108612
LGP + Pyramid	0.102825
NABP	**0.046318**
DTCTH	0.112134
MCCT	0.927301
LQCT	0.469733
mLQCT	**0.995431**

Face Recognition: LFW View2

Features	Accuracy (%)	
	Original	Aligned
LBP	67.92	69.90
LTP	68.62	72.06
LGP	63.83	67.57
HOG	67.32	69.23
Gabor [3]	62.93	-
LDN [10]	69.08	-
NABP (Proposed)	**72.58**	**74.81**

Expression Recognition: CK+

Techniques	Accuracy
SPTS [101]	50.40
CAPP [101]	66.70
SPTS+CAPP [101]	83.30
LDN [54]	89.30
LBP	88.67
LTP	89.65
LGP	83.10
HOG	89.69
CENTRIST	88.70±4.37
Proposed (NABP + Adaboost [17])	92.17
Proposed (DTCTH + Linear SVM)	93.99±5.83
Proposed (DTCTH + HI)	93.82±5.52
Proposed (MCCT + Linear SVM)	**94.27±4.57**
Proposed (MCCT + HI)	94.74±4.21
Proposed (LQCT + Linear SVM)	95.37±4.06
Proposed (LQCT + HI)	**95.55±3.48**

Gender Classification: LFW

Technique	Accuracy
Boosted LBP [51]	94.83
LBP	90.23
LTP	90.78
LGP	89.36
HOG	89.23
CENTRIST	91.92±0.34
mCENTRIST	93.19±0.60
Proposed (NABP + Adaboost [17])	92.74
Proposed (DTCTH + Linear SVM)	92.59±0.57
Proposed (DTCTH + HI)	**92.91±0.63**
Proposed (MCCT + Linear SVM)	94.35±0.43
Proposed (MCCT + HI)	**94.93±0.52**
Proposed (LQCT + Linear SVM)	93.78±0.51
Proposed (LQCT + HI)	94.69±0.57

Object Classification: Caltech 101

Training Images	5	10	15	20	25	30
SIFT [10, 33]	-	-	62.48	-	-	69.89
CS-LBP [10, 44]	-	-	58.5	-	-	66.86
DAISY [10, 72]	-	-	58.63	-	-	67.01
HSOG [10]	-	-	60.46	-	-	67.97
SVM-KNN [90]	46.6	55.8	59.05	62	-	66.23
SPM [6]	-	-	56.40	-	-	64.60
Griffin et al. [12]	44.2	54.5	59.0	63.3	65.8	67.60
NBNN [79]	-	-	65.00	-	-	70.4
ML+CORR [115]	-	-	61.00	-	-	69.60
KC [13]	-	-	-	-	-	64.14
LSPM [80]	-	-	53.23	-	-	58.81
ScSPM [80]	-	-	67.0	-	-	73.2
LLC [81]	51.15	59.77	65.43	67.74	70.16	73.44
LSA [88]	-	-	-	-	-	74.21
SP-pLSA [37]	-	-	59.8	-	-	67.7
LDC [84]	-	-	-	-	-	74.47
LCSR [82]	-	-	-	-	-	73.23
SSC [83]	55.64	65.52	69.98	73.99	75.49	77.59
PmSVM-χ^2 [91]	-	-	72.08	-	-	-
PmSVM-HI [91]	-	-	**72.18**	-	-	-
LTP + Pyramid	41.04	51.23	59.69	61.17	64.57	67.85
LGP + Pyramid	39.86	50.11	57.84	60.03	62.96	66.52
GIST	40.16	47.87	52.5	56.25	58.88	61.70
CENTRIST	39.46	49.72	55.84	59.47	62.25	65.23
Proposed (DTCTH + Linear SVM)	46.98	57.00	63.66	65.83	68.69	72.26
Proposed (DTCTH + HI)	56.74	65.97	71.84	74.80	76.85	78.56
Proposed (LQCT + Linear SVM)	49.21	60.04	65.95	69.46	72.03	74.48
Proposed (LQCT + HI)	**59.33**	**68.43**	73.48	**75.91**	**78.01**	**80.63**

Object Classification: Caltech 256

Training Images	15	30	45	50	60
SIFT [33, 94]	-	-	-	29.4	-
HOG [94]	-	-	-	33.3	-
HOG + Pyramid [94]	-	-	-	32.7	-
LBP [94]	-	-	-	20.7	-
LBP + Pyramid [94]	-	-	-	20.5	-
SPM [6]	-	34.10	-	-	-
LSPM [80]	13.20±0.62	15.45±0.37	16.37± 0.47	-	16.57±1.01
KSRSPM [85]	29.77±0.14	35.67±0.10	38.61±0.19	-	40.30±0.22
KC [13]	-	27.17±0.46	-	-	-
EMK [87]	23.2±0.6	30.5±0.4	34.4±0.4		37.6±0.5
NBNN [79]	30.4	36.0	-	-	-
Griffin et al. [12]	28.30	34.10	-	-	-
ScSPM [80]	27.73±0.51	34.02±0.35	37.46±0.55	-	40.14±0.91
LLC [81]	**34.36**	**41.19**	**45.31**	-	**47.68**
LSA [88]	-	-	-	-	36.52±0.26
LDC [84]	-	-	-	-	38.25±0.08
LScSPM [86]	29.99±0.15	35.74±0.10	38.47±0.51	-	40.32±0.32
SSC [83]	30.59±0.35	37.08±0.36	40.68±0.16	-	43.48±0.38
LTP + Pyramid	23.12±0.26	29.33±0.27	31.74±0.35	32.95±0.37	33.97±0.43
LGP + Pyramid	22.86±0.41	28.89±0.33	31.13±0.28	32.02±0.29	33.14±0.51
GIST	18.58±0.27	21.36±0.15	24.17±0.12	26.14±0.29	27.09±0.5
CENTRIST	21±0.34	27.13±0.29	29.97±0.31	31.12±0.43	32.72±0.82
Proposed (DTCTH + Linear SVM)	27.43±0.37	33.57±0.43	36.38±0.33	37.59±0.35	38.30±0.31
Proposed (DTCTH + HI)	32.91±0.31	39.42±0.21	43.07±0.18	**44.16±0.25**	45.61±0.27
Proposed (LQCT + Linear SVM)	29.97±0.32	36.12±0.38	38.75±0.28	39.83±0.37	40.62±0.29
Proposed (LQCT + HI)	35.22±0.28	41.87±0.24	45.57±0.23	**46.81±0.23**	48.93±0.30

Scene Classification: Indoor 67

Methods	Accuracy
Object Bank [75]	37.60
SIFT [33], [116]	45.86
HOG [76]	22.8
SPM [6], [9],	34.4
MM-scene [117]	28.00
DPM [76]	30.40
LSA [88]	44.19
LLC [81]	43.78
LDC [84]	46.69
PmSVM-HI [91]	**47.15**
PmSVM-χ^2 [91]	46.20
PRICoLBP [9]	43.4
SSC [83]	44.35
mSIFT [7]	39.7±1.6
mGIST [7]	31.5±1.6
LTP + Pyramid	35.87±1.23
LGP + Pyramid	34.24±1.12
GIST	26.5±1.41
CENTRIST	35.12±0.99
mCENTRIST	43.22±1.2
Proposed (DTCTH + Linear SVM)	43.33±0.72
Proposed (DTCTH + HI)	46.22±1.02
Proposed (LQCT + Linear SVM)	43.87±1.68
Proposed (LQCT + HI)	46.42±2.13
Proposed (MCCT + Linear SVM)	50.08±1.43
Proposed (MCCT + HI)	53.24±0.77
Proposed (mLQCT + Linear SVM)	50.15±0.72
Proposed (mLQCT + HI)	**53.36±1.02**

Scene Classification: OT Scene

Techniques	Accuracy
SIFT [10, 33]	84.1
HOG [10]	82.4
DAISY [10, 72]	85.7
CS-LBP [10, 44]	83.4
HSOG [10]	86.3
pLSA [93]	86.65
SPM [37]	87.1
SP-pLSA [37]	87.8
GIST	69.03
CENTRIST	84.01
LTP + Pyramid	85.6
LGP + Pyramid	84.52
mCENTRIST	87.56±0.32
Proposed (DTCTH + Linear SVM)	87.88±0.51
Proposed (DTCTH + HI)	89.18±0.81
Proposed (MCCT + Linear SVM)	89.55±0.42
Proposed (MCCT + HI)	90.50±0.59
Proposed (LQCT + Linear SVM)	88.97±0.73
Proposed (LQCT + HI)	89.89±0.80
Proposed (mLQCT + Linear SVM)	89.66±0.52
Proposed (mLQCT + HI)	**90.81±0.56**

Scene Classification: Scene 15

Methods	Accuracy
SPM [6]	81.40±0.50
Object Bank [75]	80.90
SIFT [33, 116]	82.06
ScSPM [80]	80.28±0.93
SC + linear kernel [89]	**84.10±0.50**
NBNN [78, 79]	72.3±0.93
I2CDML [78]	77.00±0.6
I2CDML+SPM [78]	81.2±0.52
LLC [81, 84]	79.81±0.35
LSA [88]	80.12±0.60
pLSA [37]	72.7
SP-pLSA [37]	83.7
SPCK++ [77]	82.51±0.43
LDC [84]	82.50±0.47
LCSR [82]	82.67±0.51
PRICoLBP [9]	82.04
LTP + Pyramid	80.25±0.31
LGP + Pyramid	78.22±0.56
GIST	55.55±0.67
CENTRIST	81.45±0.23
Proposed (DTCTH + Linear SVM)	82.66±0.5
Proposed (DTCTH + HI)	83.63±0.21
Proposed (LQCT + Linear SVM)	86.13±0.68
Proposed (LQCT + HI)	**87.01±0.42**

Event Classification: UIUC Sports Event

Techniques	Accuracy
KSRSPM [85]	84.92±0.78
ScSPM [80]	82.74
SIFT [33, 116]	85.12
LSA [88]	82.29±1.84
LLC [81, 88]	81.41±1.84
LCSR [82]	87.23±1.14
NBNN [78, 79]	67.6±1.1
I2CDML [78]	78.5±1.63
I2CDML+SPM [78]	79.7±1.83
LQP [43, 74]	78.9
DDLBP + Max Relevance [43]	83.5
DDLBP + mRMR [43]	83.5
DDLBP + MJMI [43]	84.0
mGIST [7]	76.2±1.9
mSIFT [7]	84.2±0.7
mCENTRIST [7]	86.5±0.6
LTP + Pyramid	82.43±1.17
LGP + Pyramid	78.42±0.94
GIST	69.95±0.98
CENTRIST	79.50±0.95
mCENTRIST	85.58±1.91
Proposed (DTCTH + Linear SVM)	85.16±0.96
Proposed (DTCTH + HI)	88.18±0.84
Proposed (MCCT + Linear SVM)	88.01±1.15
Proposed (MCCT + HI)	**90.13±0.32**
Proposed (LQCT + Linear SVM)	86.88±0.72
Proposed (LQCT + HI)	88.89±0.73
Proposed (mLQCT + Linear SVM)	**88.36±0.78**
Proposed (mLQCT + HI)	89.58±0.59

Flower Classification: Oxford Flower 102

Technique	Accuracy
SIFT_B [106]	32.0
HOG [106]	49.6
SIFT internal [106]	55.1
KMTJSRC-CG (SIFTint) [95]	55.2
Yuan et al. [95]	71.2
LLC [81, 84]	57.75
LSA [84, 88]	57.8
LDC [84]	61.45
PRICoLBP + **Segmentation** [9]	82.3
LTP + Pyramid	51.18±0.42
LGP + Pyramid	43.82±0.55
GIST	23.22±0.50
CENTRIST	49.28±0.40
mCENTRIST	68.58±1.91
Proposed (DTCTH + Linear SVM)	53.43±0.81
Proposed (DTCTH + HI)	57.71±0.47
Proposed (MCCT + Linear SVM)	72.89±0.70
Proposed (MCCT + HI)	78.85±0.32
Proposed (LQCT + Linear SVM)	64.94±0.32
Proposed (LQCT + HI)	70.90±0.54
Proposed (mLQCT + Linear SVM)	**76.87±0.28**
Proposed (mLQCT + HI)	**81.51±0.47**

Aerial Image Classification: Land-Use 21

Techniques	Land-Use 21
SPCK++ [77]	77.38
LQP [43, 74]	83.0
DDLBP + Max Relevance [43]	86.3
DDLBP + mRMR [43]	87.0
DDLBP + MJMI [43]	87.2
mSIFT [7]	85.0±2.6
mGIST [7]	72.0±2.7
LTP + Pyramid	80.33±1.72
LGP + Pyramid	75.57±2.67
GIST	52.57±2.78
CENTRIST	78.10±1.52
mCENTRIST	89.62±2.13
Proposed (DTCTH + Linear SVM)	82.57±0.49
Proposed (DTCTH + HI)	85.89±1.57
Proposed (MCCT + Linear SVM)	91.60±0.46
Proposed (MCCT + HI)	92.36±1.16
Proposed (LQCT + Linear SVM)	88.62±1.64
Proposed (LQCT + HI)	88.86±1.52
Proposed (mLQCT + Linear SVM)	91.38±1.62
Proposed (mLQCT + HI)	**92.60±0.92**

Scene Classification: SUN 397

Methods	Accuracy
Xiao et al. [103]	27.50
Kwitt et al. [118]	28.90
Lu et al. [119]	30.50
denseSIFT [94]	21.50
sparseSIFT [94]	11.50
HOG [94]	27.20
LBP [94]	18.0
LTP + Pyramid	21.25±0.31
LGP + Pyramid	18.22±0.25
GIST	16.30±0.21
CENTRIST	19.35±0.26
mCENTRIST	27.21±0.29
Proposed (DTCTH + Linear SVM)	24.87±0.23
Proposed (DTCTH + HI)	27.32±0.28
Proposed (LQCT + Linear SVM)	29.13±0.26
Proposed (LQCT + HI)	33.27±0.29
Proposed (MCCT + Linear SVM)	32.14±0.53
Proposed (MCCT + HI)	38.24±0.46
Proposed (mLQCT + Linear SVM)	33.76±0.47
Proposed (mLQCT + HI)	39.53±0.41

Conclusion and Future Work

Conclusion

- MCCT performs well in color image classification
- LQCT performs well in gray-scale image while its extension mLQCT performs well in color image classification
- mLQCT > LQCT > MCCT > DTCTH > NABP (with respect to response time)

Future Work

- High level representation will be incorporated with all the proposed descriptors
  - Sparse coding
  - Pooling techniques

Publications

1. “Noise adaptive binary pattern for face image analysis," in Computer and Information Technology (ICCIT), 2015 18th International Conference on. IEEE, 2015. The contribution of chapter 5. (**1st prize, best paper award, 18th ICCIT, 2015**)

2. “DTCTH: A Discriminative Local Pattern Descriptor for Image Classification,“ Eurasip Journal of Image and Video Processing, March, 2016. The contribution of chapter 6. (Submitted)

3. “MCCT: A Multi-channel Complementary Census Transform for Image Classification," Journal of Signal, Image and Video Processing, April, 2016. The contribution of chapter 7. (Submitted)

References

[1] Déniz, Oscar, Gloria Bueno, Jesús Salido, and Fernando De la Torre. "Face recognition using histograms of oriented gradients." Pattern Recognition Letters 32, no. 12 (2011): 1598-1603.

[2] Lowe, David G. "Object recognition from local scale-invariant features." In Computer vision, 1999. The proceedings of the seventh IEEE international conference on, vol. 2, pp. 1150-1157. IEEE, 1999.

[3] Yang, Meng, Lei Zhang, Simon CK Shiu, and David Zhang. "Gabor feature based robust representation and classification for face recognition with Gabor occlusion dictionary." Pattern Recognition 46, no. 7 (2013): 1865-1878.

[4] Ojala, Timo, Matti Pietikäinen, and Topi Mäenpää. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." Pattern Analysis and Machine Intelligence, IEEE Transactions on 24, no. 7 (2002): 971-987.

[5] Ahonen, Timo, Abdenour Hadid, and Matti Pietikainen. "Face description with local binary patterns: Application to face recognition." Pattern Analysis and Machine Intelligence, IEEE Transactions on 28, no. 12 (2006): 2037-2041.

[6] Jun, Bongjin, Inho Choi, and Daijin Kim. "Local transform features and hybridization for accurate face and human detection." Pattern Analysis and Machine Intelligence, IEEE Transactions on 35, no. 6 (2013): 1423-1436.

[7] Tan, Xiaoyang, and Bill Triggs. "Enhanced local texture feature sets for face recognition under difficult lighting conditions." Image Processing, IEEE Transactions on 19, no. 6 (2010): 1635-1650.

References

[8] Murala, Subrahmanyam, R. P. Maheshwari, and R. Balasubramanian. "Local tetra patterns: a new feature descriptor for content-based image retrieval." Image Processing, IEEE Transactions on 21, no. 5 (2012): 2874-2886.

[9] Ramirez Rivera, Adin, Jorge Rojas Castillo, and Oksam Chae. "Local directional number pattern for face analysis: Face and expression recognition." Image Processing, IEEE Transactions on 22, no. 5 (2013): 1740-1752.

[10] Zhang, Baochang, Yongsheng Gao, Sanqiang Zhao, and Jianzhuang Liu. "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor." Image Processing, IEEE Transactions on 19, no. 2 (2010): 533-544.

[11] Mu, Yadong, Shuicheng Yan, Yi Liu, Thomas Huang, and Bingfeng Zhou. "Discriminative local binary patterns for human detection in personal album." In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, pp. 1-8. IEEE, 2008.

[12] Lucey, Patrick, Jeffrey F. Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression." In Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, pp. 94-101. IEEE, 2010.

[13] Markus Greiner, CA Regal, JT Stewart, and DS Jin. Probing pair-correlated fermionic atoms through correlations in atom shot noise. Physical review letters, 94(11):110401, 2005.

Any Question?

IIT
University of Dhaka

Thank You

Appendix

Face Recognition

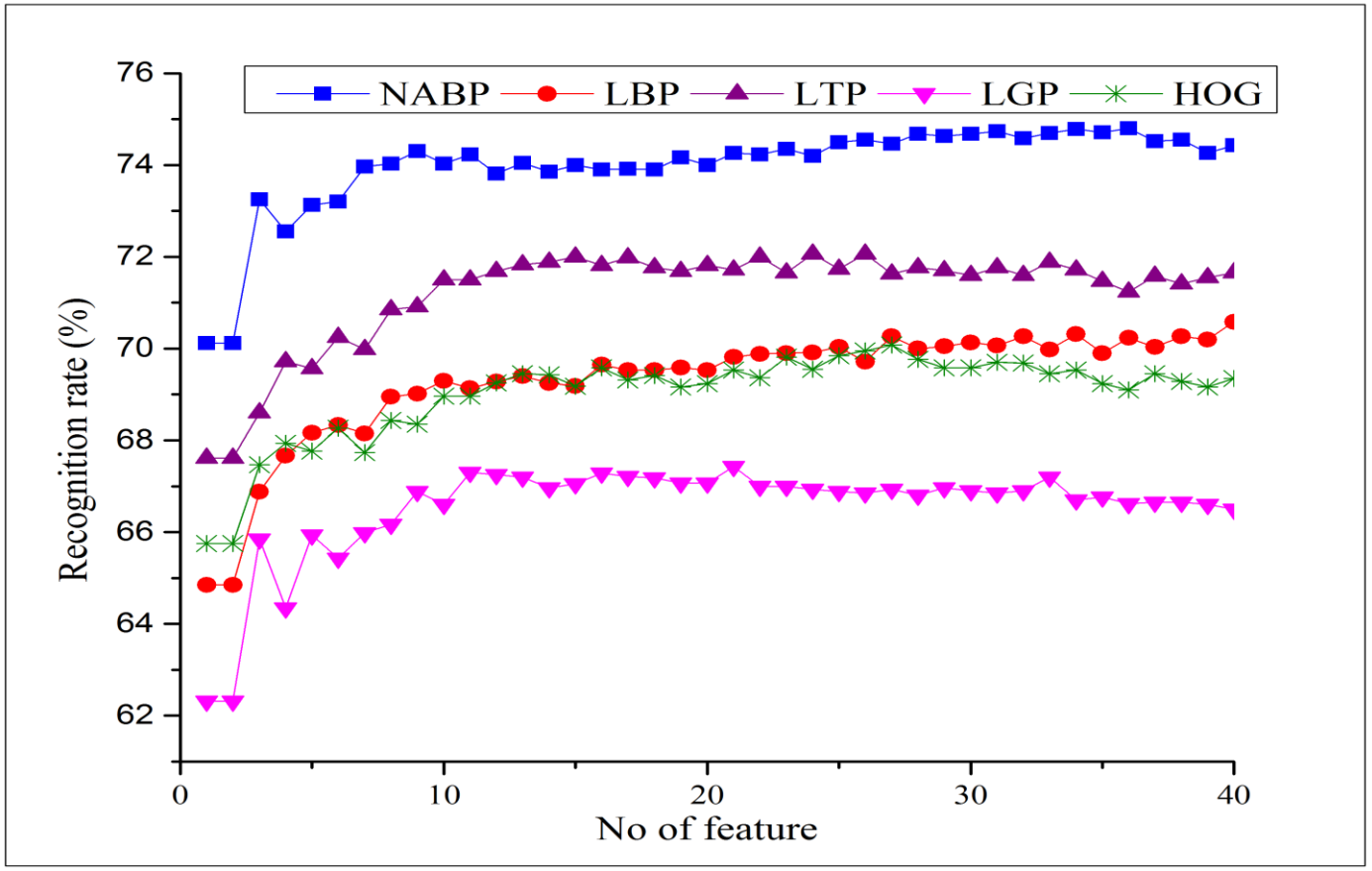


Fig. 36: Face recognition results in LFW

Expression Recognition: CK

Techniques	CK	
	6-class expression	7-class expression
LBP [14]	92.6 ± 2.9	88.9 ± 3.5
LBP + Template Matching [18]	84.5 ± 5.2	79.1 ± 4.6
Geometric Feature + TAN [113]	-	73.2
LBP + SVM [18]	91.5±3.1	88.1±3.8
Boosted-LBP [18]	89.8 ± 4.7	85.0 ±4.5
Boosted-LBP + SVM [18]	95.0 ±3.2	91.1 ± 4.0
Gabor + SVM [114]	-	84.8
Gabor [18]	89.4 ± 3.0	86.6 ± 4.1
LDN [54]	98.4±1.4	92.3±3.0
LTP + Pyramid	91.18±8.68	88.79±2.31
LGP + Pyramid	93.36±3.76	88.97±4.18
CENTRIST	89.84±7.90	86.69±2.04
Proposed (DTCTH + Linear SVM)	**98.98±1.29**	92.75±5.43
Proposed (DTCTH + HI)	97.76±2.43	**93.89±2.63**
Proposed (MCCT + Linear SVM)	99.52	Color image
Proposed (MCCT + HI)	**99.61**	Color image

Expression Recognition: CK

Table II: Confusion matrix of DTCTH for 6 class expression recognition rate in CK

	Anger	Disgust	Fear	Sadness	Happy	Surprise
Anger	**99.22**	0.0	0.78	0.0	0.0	0.0
Disgust	0.0	**100.0**	0.0	0.0	0.0	0.0
Fear	0.0	0.0	**97.22**	0.0	2.78	0.0
Sadness	0.83	0.0	0.0	**98.33**	0.0	0.83
Happy	0.43	0.0	0.85	0.0	**98.72**	0.0
Surprise	0.0	0.0	0.0	0.0	0.0	**100.0**

Expression Recognition: CK

Table II: Confusion matrix of DTCTH for 7 class expression recognition rate in CK

	Anger	Disgust	Fear	Sadness	Happy	Neutral	Surprise
Anger	**86.67**	0.0	1.90	1.9	0.0	9.52	0.0
Disgust	0.77	**95.38**	0.0	0.0	0.0	3.85	0.0
Fear	0.56	0.0	**95.0**	0.0	0.56	3.89	0.0
Sadness	1.67	0.0	0.0	**93.9**	0.0	3.89	0.56
Happy	0.42	0.0	0.0	0.0	**99.2**	0.42	0.0
Neutral	2.71	0.21	1.67	0.21	0.63	**94.58**	0.0
Surprise	0.0	0.0	0.0	0.0	0.0	0.0	**100**

Expression Recognition: CK+

Table II: Confusion matrix of DTCTH for 7 class expression recognition rate in CK+

	Anger	Contempt	Disgust	Fear	Sadness	Happy	Surprise
Anger	**96.3**	2.22	0.74	0.0	0.74	0.0	0.0
Contempt	9.26	**87.04**	0.0	0.0	0.0	0.0	0.0
Disgust	0.56	0.0	**99.44**	0.0	0.0	0.0	0.0
Fear	0.0	0.0	1.33	**90.67**	0.0	8.0	0.0
Sadness	11.9	1.19	0.0	0.0	**85.7**	0.0	1.19
Happy	0.0	0.0	0.0	0.97	0.0	**99.03**	0.0
Surprise	0.0	0.0	0.40	0.0	0.0	0.0	**99.6**

Leaf Classification: Swedish Leaf

Techniques	Accuracy	Input
Soderkvist [97]	82.40	Contour only
SC + DP [98]	88.12	Contour only
CENTRIST [8, 9]	90.61	Contour only
IDSC + DP [98]	94.13	Contour only
SPTC + DP [98]	95.33	Gray-scale image
Shape-Tree [96]	96.28	Contour only
SLPA [99]	96.33	Gray-scale image
PRICoLBP [9]	99.38	Color image
LTP + Pyramid	98.20	Gray-scale image
LGP + Pyramid	98.08	Gray-scale image
GIST	96.08	Gray-scale image
CENTRIST	97.44	Gray-scale image
mCENTRIST	99.39	Color image
Proposed (DTCTH + Linear SVM)	99.49	Gray-scale image
Proposed (DTCTH + HI)	**99.52**	Gray-scale image
Proposed (MCCT + Linear SVM)	99.52	Color image
Proposed (MCCT + HI)	**99.61**	Color image
Proposed (LQCT + Linear SVM)	99.57±0.09	Gray-scale image
Proposed (LQCT + HI)	99.68±0.07	Gray-scale image
Proposed (mLQCT + Linear SVM)	99.61±0.09	Color image
Proposed (mLQCT + HI)	**99.72±0.06**	Color image

Scene Classification: RGB NIR Scene 9

Methods	Accuracy
LTP + Pyramid	74.80±4.39
LGP + Pyramid	71.54±5.48
GIST	72.12±5.64
CENTRIST	75.60±4.16
mCENTRIST	81.53±3.94
Proposed (DTCTH + Linear SVM)	79.72±5.31
Proposed (DTCTH + HI)	81.22±4.97
Proposed (LQCT + Linear SVM)	80.39±3.08
Proposed (LQCT + HI)	81.59±4.18
Proposed (MCCT + Linear SVM)	85.71±4.78
Proposed (MCCT + HI)	86.71±4.04
Proposed (mLQCT + Linear SVM)	**86.75±3.21**
Proposed (mLQCT + HI)	85.92±5.10

Garments Texture Classification: Fashion

Techniques	5 class	3 class
LTP + Pyramid	79.64±0.78	77.62±0.72
LGP + Pyramid	78.57±0.67	76.23±0.81
GIST	62.59±1.48	60.85±0.63
CENTRIST	78.89±0.81	75.23±0.57
mCENTRIST	82.62±0.49	80.21±0.86
Proposed (DTCTH + Linear SVM)	85.29±0.59	82.33±0.76
Proposed (DTCTH + HI)	**85.97±0.61**	82.65±0.78
Proposed (MCCT + Linear SVM)	87.60±0.46	85.31±0.43
Proposed (MCCT + HI)	**88.36±0.52**	86.46±0.49
Proposed (LQCT + Linear SVM)	86.53±0.23	85.56±0.73
Proposed (LQCT + HI)	**87.5±0.45**	86.92±0.45
Proposed (mLQCT + Linear SVM)	88.52±0.62	87.33±0.72
Proposed (mLQCT + HI)	**89.97±0.47**	88.33±0.72

Garments Texture Classification: Clothing Attribute

Techniques	Accuracy
LTP + Pyramid	75.25±1.34
LGP + Pyramid	73.53±1.86
GIST	58.98±2.23
CENTRIST	74.17±1.26
mCENTRIST	80.19±1.16
Proposed (LQCT + Linear SVM)	84.56±0.87
Proposed (LQCT + HI)	82.53±2.35
Proposed (mLQCT + Linear SVM)	81.60±2.10
Proposed (mLQCT + HI)	**86.29±1.27**

Texture Classification: FMD

Techniques	Accuracy
LTP + Pyramid	34.33±1.72
LGP + Pyramid	33.57±1.36
GIST	23.79±1.26
CENTRIST	35.10±1.52
mCENTRIST	89.62±2.13
Proposed (DTCTH + Linear SVM)	46.28±0.46
Proposed (DTCTH + HI)	48.88±1.16
Proposed (LQCT + Linear SVM)	45.20±1.12
Proposed (LQCT + HI)	49.48±1.31
Proposed (mLQCT + Linear SVM)	49.47±0.82
Proposed (mLQCT + HI)	**54.00±1.36**

Texture Classification: K'th Tips

Techniques	Accuracy
LTP + Pyramid	88.35±0.75
LGP + Pyramid	87.16±0.86
GIST	68.64±1.45
CENTRIST	88.13±0.75
mCENTRIST	95.87±0.54
Proposed (MCCT + Linear SVM)	98.15±0.49
Proposed (MCCT + HI)	**98.46±0.67**
Proposed (LQCT + Linear SVM)	98.39±0.56
Proposed (LQCT + HI)	**98.74±0.66**
Proposed (mLQCT + Linear SVM)	98.74±0.43
Proposed (mLQCT + HI)	**99.03±0.31**

Skin Disease Classification: 3-Skin Disease

Techniques	Accuracy
LTP + Pyramid	71.23±1.16
LGP + Pyramid	70.25±1.67
GIST	47.51±2.41
CENTRIST	69.13±1.27
mCENTRIST	79.52±1.13
Proposed (MCCT + Linear SVM)	84.60±1.46
Proposed (MCCT + HI)	88.36±1.28
Proposed (LQCT + Linear SVM)	84.65±1.34
Proposed (LQCT + HI)	85.24±1.63
Proposed (mLQCT + Linear SVM)	88.06±1.99
Proposed (mLQCT + HI)	**89.35±0.96**