University of Nottingham
UK | CHINA | MALAYSIA

icassp 2022 Singapore

Paper ID: 5984

Dynamic Texture Recognition using PDV Hashing and Dictionary Learning on Multi-scale Volume Local Binary Pattern

Ruxin Ding, Jianfeng Ren, Heng Yu, Jiawei Li

School of Computer Science
University of Nottingham Ningbo China
199 Taikang East Road, Ningbo, 315100 China

University of Nottingham
UK | CHINA | MALAYSIA

Contents

University of Nottingham
UK | CHINA | MALAYSIA

Introduction

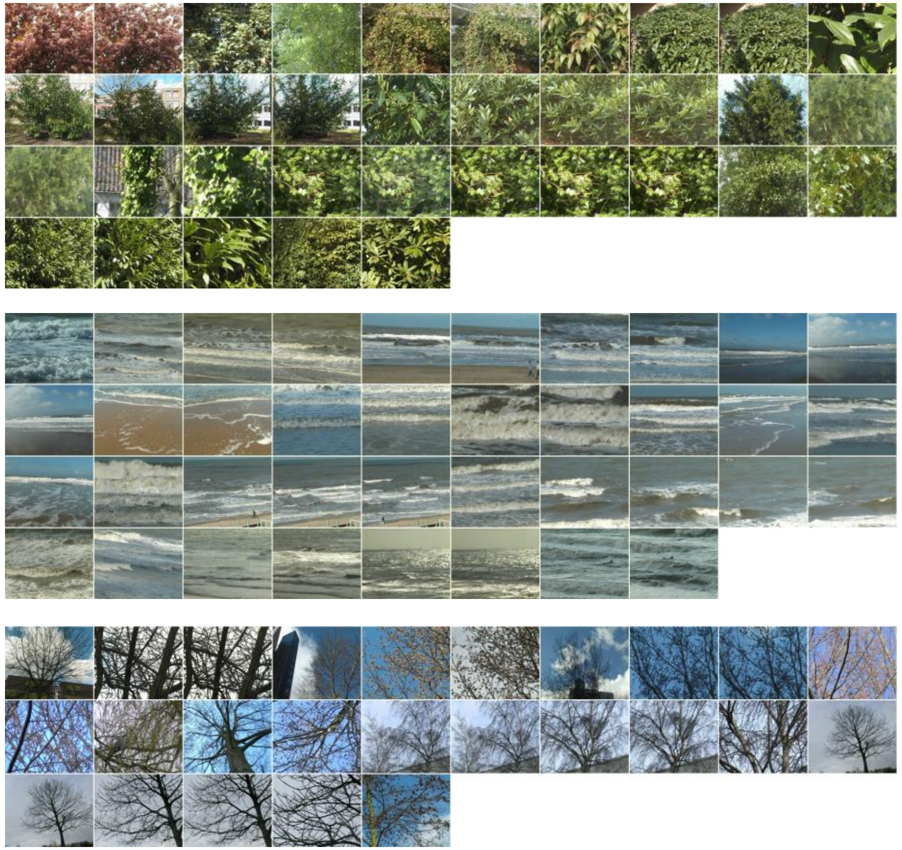
Dynamic Texture Recognition

Dynamic textures (DTs) refer to sequences of the image that consists of repeated patterns related to time and space.

Applications: facial analysis, video retrieval, lip reading, micro expression analysis, etc.

The methods for DT recognition can be based on:

1) generative model.
2) optical flow.
3) geometry.
4) 3D filters.
5) **local features.**
6) learning.



Example DTs from [1]

[1] Renaud Peteri, S ´ andor Fazekas, and Mark J. Huiskes, “Dyn- ´ tex: A comprehensive database of dynamic textures,” Pattern Recognition Letters, vol. 31, no. 12, pp. 1627–1632, 2010.

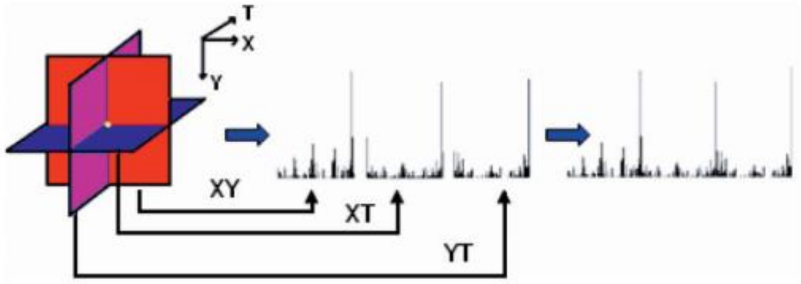
Introduction

Volume Local Binary Pattern (VLBP) [2]

+ Combines the motion and appearance features together
- The feature dimension of VLBP increases exponentially with the number of neighbors

Local Binary Patterns From Three Orthogonal Planes (LBP-TOP) [2]

+ Combines the motion and appearance features together
+ Reduce the dimension of VLBP
- The correlation information among the three planes is lost



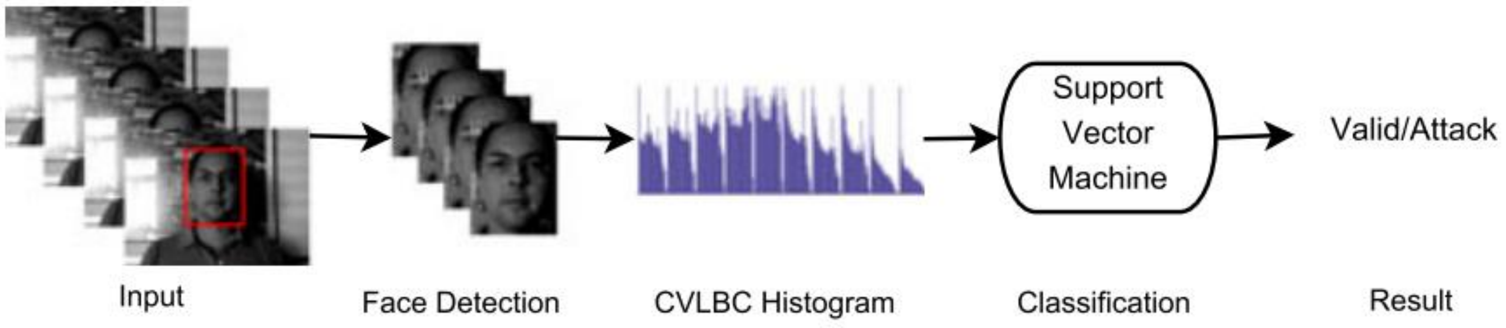
(a) (b) (c)

[2] G. Zhao and M. Pietikainen, “Dynamic texture recognition using local binary patterns with an application to facial expressions,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 915–928, 2007.

Introduction

Completed Volume Local Binary Count (CVLBC)[3] ← VLBC[3] ↑ LBC[4]

+ Both the local difference and volume center pixel intensity are utilized
+ Extracting the magnitude of local differences & information of central pixels and encoding them into a compact code
- The dimensional problem can still be improved

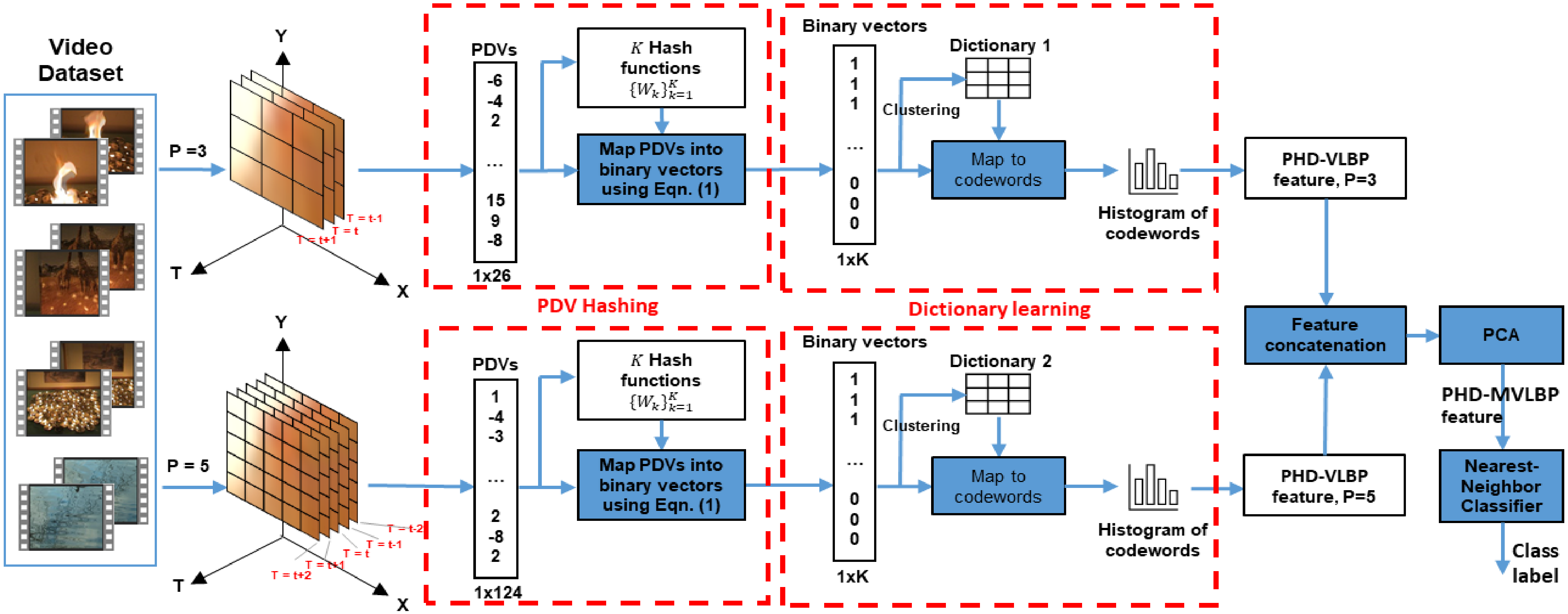


[3]X. Zhao, Y. Lin, and J. Heikkil ̈a, "Dynamic texture recognition using volume local binary count patterns with an application to 2D face spoofing detection," IEEE Transactions on Multimedia, vol. 20, no. 3, pp. 552–566, 201
[4] Y. Zhao, D.-S. Huang, and W. Jia, "Completed local binary count for rotation invariant texture classification," IEEE Trans. Image Process., vol. 21, no. 10, pp. 4492–4497, Oct. 2012.

Contents

- Introduction
- **Proposed Method**
- Experimental Results
- Conclusion

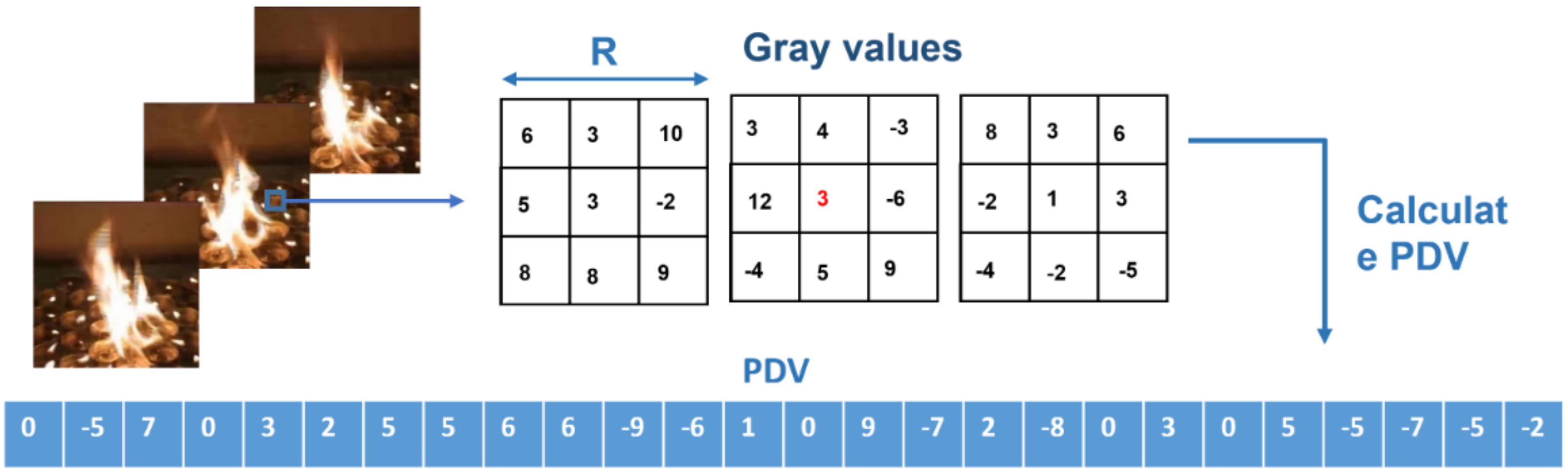
PDV hashing and dictionary learning on multi-scale VLBP (PHD-MVLBP)



PDV hashing and dictionary learning on multi-scale VLBP (PHD-MVLBP)

Obtain Pixel Difference Vectors (PDVs)

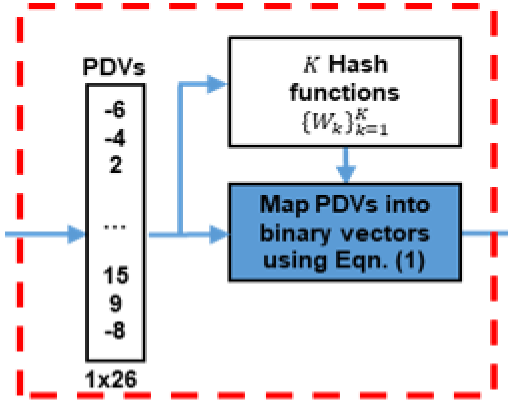
Give the neighborhood of size $P \times P \times P$, the center pixel is I_c and the neighboring pixels are $I_1, I_2, \ldots, I_{P^3-1}$, respectively, the pixel difference vector $x = [I_1, I_2, \ldots, I_{P^3-1}] - I_c \in R^{P^3-1}$.



PDV hashing and dictionary learning on multi-scale VLBP (PHD-MVLBP)

Mapping PDVs into Binary Vectors Using Hash Functions

Inspired by [5], K hash functions are used to map each x_n into a binary vector $b_n = [b_{1n}, \dots, b_{kn}]^T \in \{0,1\}^{K\times 1}$.



PDV Hashing

$$b_{kn} = 0.5 \times \left(sgn\left(w_k^T x_n\right) + 1\right)$$

$$\min_{w_k} J = J_1 + \lambda_1 J_2 + \lambda_2 J_3 + \lambda_3 J_4$$

$$= \sum_{n=1}^{N} \left\| \sum_{k=1}^{K-1} \left\| b_{kn} - b_{(k+1)n} \right\|^2 - 1 \right\|^2$$

$$+\lambda_1 \sum_{n=1}^{N} \sum_{k=1}^{K} \left\| (b_{kn} - 0.5) - w_k^T x_n \right\|^2$$

$$+\lambda_2 \sum_{k=1}^{K} \left\| \sum_{n=1}^{N} (b_{kn} - 0.5) \right\|^2$$

$$-\lambda_3 \sum_{n=1}^{N} \sum_{k=1}^{K} \left\| b_{kn} - \mu_k \right\|^2$$

J1: make the adjacent bits as equal as possible & prevent all zeros or ones

J2: minimize the loss of energy

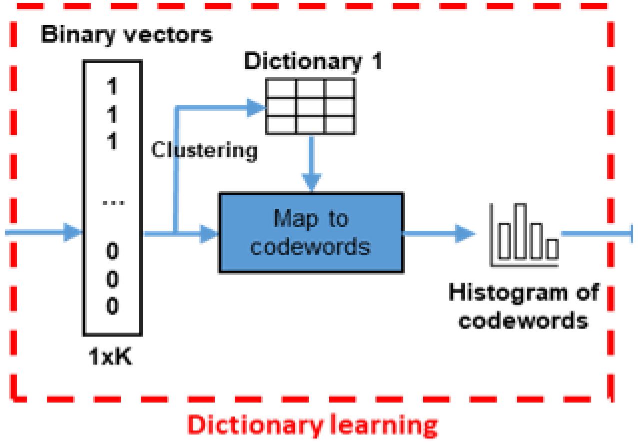
J3: to evenly distribute each feature bit

J4: maximize the variance of binary codes

[5] Y. Duan, J. Lu, J. Feng, and J. Zhou, "Context-aware local binary feature learning for face recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 5, pp. 1139–1153, 2018.

PDV hashing and dictionary learning on multi-scale VLBP (PHD-MVLBP)

Dictionary Learning for Binary Vectors

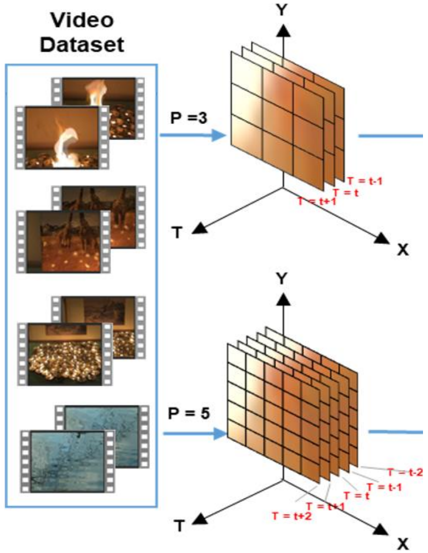


The binary vectors are clustered to form the dictionary of D codewords.

PDVs are mapped to binary vectors using hash functions -> mapped to histogram of codewords using the derived dictionary.

Multi-scale Feature Extraction

Features extracted at different scales *P* are concatenated.
In our method, we take *P = 3* and *P = 5*.



Contents

- Introduction
- Proposed Method
- **Experimental Results**
- Conclusion

Experimental Results

DynTex++ dataset

- 36 classes, each contains 100 videos (50 × 50 × 50)
- 50% for training and the other for testing
- Repeated 5 times
- The dictionary size is 1500, PCA dimension is 500
- Nearest-neighbor classifier with cosine distance

Method	Accuracy
Distance Learning	63.70%
VLBP	87.35%
LBP-TOP	93.20%
DFA	89.90%
CVLBC	91.31%
MBSIF-TOP	97.17%
Proposed PHD-VLBP, *P=3*	**97.51%**
Proposed PHD-VLBP, *P=5*	97.10%
Proposed PHD-MVLBP	**97.77%**

Experimental Results

UCLA dataset

- 200 DT videos in total, including 50 scenes, with 4 sequences for each scene (75 × 48 × 48)
- Repeated 20 times
- Nearest-neighbor classifier with cosine distance

- 50-Class Breakdown: 4-fold cross-validation
- 9-Class Breakdown: 50% for training and the other for testing

Method	50-class	9-class
Distance Learning	99.00%	95.60%
VLBP	-	96.30%
LBP-TOP	-	96.00%
KDT-MD	97.50%	97.50%
CVLBC	99.50%	-
MBSIF-TOP	99.50%	98.75%
Proposed PHD-VLBP, *P=3*	**100.00%**	98.65%
Proposed PHD-VLBP, *P=5*	**100.00%**	98.50%
Proposed PHD-MVLBP	**100.00%**	**98.90%**

Contents

Conclusion

- PDV hashing and dictionary learning on multi-scale VLBP (PHD-MVLBP)
- Tackle the problems of high dimensionality of VLBP and the potential information loss of LBP-TOP
- Effectively extract spatial-temporal discriminant information from videos
- Demonstrates superior performance compared with the state-of-the-art methods on Dyntex++ and UCLA datasets

THANK YOU