CHARACTERIZING DYNAMIC TEXTURES WITH SPACE-TIME LACUNARITY ANALYSIS

Yuping Sun[a,b], Yong Xu[a] and Yuhui Quan[b]*

[a]School of Computer Science & Engineering, South China University of Technology, Guangzhou, China
[b]Department of Mathematics, National University of Singapore, Singapore
sun.yp@mail.scut.edu.cn yxu@scut.edu.cn matquan@nus.edu.sg

ABSTRACT

This paper addresses the challenge of reliably capturing the temporal characteristics of local space-time patterns in dynamic texture (DT). A powerful DT descriptor is proposed, which enjoys strong robustness to viewpoint changes, illumination changes, and video deformation. Observing that local DT patterns are spatial-temporally distributed with stationary irregularities, we proposed to characterize the distributions of local binarized DT patterns along both the temporal and the spatial axes via lacunarity analysis. We also observed such irregularities are similar on the DT slices along the same axis but distinct between axes. Thus, the resulting lacunarity based features are averaged along each axis and concatenated as the final DT descriptor. We applied the proposed DT descriptor to DT classification and evaluated its performance on several benchmark datasets. The experimental results have demonstrated the power of the proposed descriptor in comparison with existing ones.

***Index Terms*—** Dynamic Texture, Local Binarized Patterns, Stationary Irregularities, Lacunarity Analysis

1. INTRODUCTION

Dynamic textures (DTs) refer to video sequences of moving scenes, which exhibit certain stationarity properties in time domain [1]. Such sequences are very common in real world, such as video clips of rivers, waterfall, foliage, smoke, clouds, and fire. See Fig. 1 for some examples of DT sequences. The potential applications concerning DT analysis are significant and numerous, including remote monitoring, video retrieval, foreground-background separation, and various types of surveillance. Compared with static texture, DT is more challenging to analyze, owing to the additional difficulties in reliably capturing the temporal characteristics of texture elements. Moreover, clutters, occlusions, video deformation, and a wide range of environmental changes like the changes of illumination and viewpoint, make the characterization of DT more difficult. This inspired us to develop an effective descriptor for DT characterization.

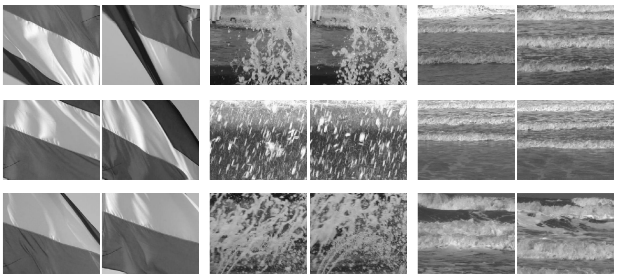

Fig. 1. Examples of DT sequences. Each column shows three DT sequences from the same category. For each DT sequence only a pair of key frames are shown.

The topics of DT characterization range from DT modeling and synthesis to DT segmentation and classification. The focus of this paper is on the development of an effective DT descriptor for robustly classifying DTs with complex motion.

1.1. Related work

There have been a variety of methods proposed for DT classification, which can be roughly classified into two groups:

- Parametric methods. A line of research is based on the so-called parametric methods [3, 4, 5, 6]. In these methods, the generation of DTs is modeled by certain underlying physical dynamic system and the estimated parameters of the corresponding model are used for DT classification. There is a broad spectrum of parametric DT models, e.g. the spatio-temporal autoregressive model [4] which expresses each DT voxel by the linear combination of it spatio-temporal neighbors, the linear dynamical system [3] with parameters analyzed on the Stiefel manifold, the Bag-of-Systems model [7] based on the distributions of local dynamical models from

Yuping Sun would like to thank the partial support by China Scholarship Council Program(CSC)([2013]3009). Yong Xu (*corresponding author), would like to thank the supports by National Nature Science Foundations of China (61273255 and 61070091), Engineering & Technology Research Center of Guangdong Province for Big Data Analysis and Processing ([2013]1589-1-11), Project of High Level Talents in Higher Institution of Guangdong Province (2013-2050205-47) and Guangdong Technological Innovation Project (2013KJCX0010). Yuhui Quan would like to thank the partial support by Singapore MOE Research Grant R-146-000-178-112.

the DT sequence, and the dynamic multi-scale autoregressive models [8, 9]. Although the parametric methods provide good understanding of the essence of DTs, they encounter the inflexibility in describing the DTs generated by nonlinear physical systems with complex motion irregularities.

- Non-parametric methods. Apart from the parametric methods, there are other approaches which extract discriminative information from DT sequences without explicitly modeling the underlying dynamic patterns by certain DT generation systems. Such approaches, referred to as non-parametric methods, can be further categorized into various types, depending on the techniques used for DT classification. One type is the field-based approaches in which the classification is done on the motion field [2, 10, 11]. Using the estimated instantaneous motion patterns in DT sequences, Chetverikov et al. [2] proposed to convert the analysis of spatio-temporal sequences to that of sequences of static information. Peteri et al. [10] proposed to extract six translation invariant features based on normal flow and texture regularity to describe the dynamics and appearance of DT sequences. A metric of video sequences is defined in [11] using the velocity and acceleration fields estimated at various spatio-temporal scales. The main drawback of these methods is their heavy dependence on the estimation of motion field within video frames, which is likely to fail in stochastic dynamics lacking of brightness constancy and local smoothness.

Alternatively, some approaches have been proposed to collect histogram-based statistics of certain local dynamic patterns [12, 13, 14]. The performance of such approaches is highly dependent on the spatio-temporal appearance of DTs captured by the local descriptors. The choices of the local descriptors vary from spatio-temporal wavelet coefficients [15, 16] to space-time oriented patterns [12, 17, 18]. Recently, the local binary pattern (LBP) [19, 20] has emerged as a simple yet powerful local descriptor due to its robustness to monotonic gray-scale changes and alignment error. Zhao et al. [13] proposed two types of histogram-based features based on the volume local binary patterns and local binary patterns from three orthogonal planes. However, even with robust local descriptors, the performance of these methods in recognizing complex DT sequences is limited as simple histogram-based statistics would lose the details of the spatio-temporal distribution of the dynamic patterns.

In order to address the aforementioned issues, most recently, the fractal-based methods [16, 21, 22] have been proposed which can be viewed as discriminative methods with generative motivations. The basic idea of these approaches is to treat a DT sequence as being generated by some physical dynamic process with spatio-temporal self-similarities in its dynamics and then apply multi-fractal spectrum to characterize such self-similarities without explicitly defining the dynamic process. However, the local descriptors used in these methods, such as intensity and gradient [21], wavelet coefficients [16] and predefined templates [22], are sensitive to gray-scale changes and might produce big quantization error in the stage of partitioning the feature bins.

Finally, it should be mentioned that some approaches have been proposed to combine the discriminative methods with the generative ones. For example, Ghanem et al. [23] combined the two discriminative descriptors with one generative descriptor by adaptive weighting based on maximum margin distance learning.

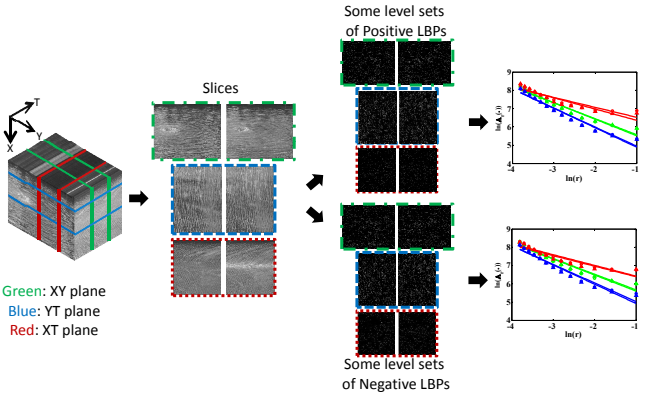

Fig. 2. Outline of the proposed method. Given a DT sequence, the 2D slices are sampled along the X, Y, and T axes. Then for each 2D slice, two types of local patterns are extracted, resulting in multiple binary images generated by pixel-wise classification. Finally, lacunarity analysis is applied to each binary image via bi-logarithmic least square fitting on the number of non-empty boxes versus box size. The resulting lacunarity features on the 2D slices are collected as the final DT feature. It is seen in the second column that the distributions of local DT patterns over the spatial and temporal domain exhibit significant difference on the 2D slices sampled from different planes while being similar on the slices along the same axis, which serves as the motivation of our method.

1.2. Motivation and contributions

A desirable DT descriptor for classification should not only be discriminative but also enjoy strong robustness to a wide range of environmental changes like the changes of viewpoint and illumination. We observed that the local space-time patterns are distributed with stationary irregularities over space and time. Besides, such irregularities along different axes (i.e., the horizontal X, vertical Y, and temporal T axes) are significantly discriminative to each other. See the 2D slices in Fig. 1 for an illustration. Thus, we are inspired to characterize DT by describing the irregularities of the spatial and temporal distributions of local space-time patterns extracted by some effective local pattern encoding schemes.

In the paper, an effective method is proposed for DT classification, which examines the spatio-temporal distribution of DT voxels from three orthogonal planes respectively. To efficiently locate various types of local DT structures on the 2D slices sampled from different planes, we employ both the

positive and negative local pattern encoding schemes, which are not only robust to monotonic gray-scale changes but also resistant to random and quantization noise. Then we apply lacunarity analysis to capture the stationary spatial and temporal irregularities of the distribution of each type of local binarized patterns along each axis, leading to highly discriminative DT features with strong robustness to a wide range of distortions and viewpoint changes. As the distributions of the local DT patterns over the space and time domain are similar on the slices along the same axis, the resulting slice-wise feature vectors are averaged along each axis respectively and then concatenated as the final DT descriptor. The experimental results have demonstrated the excellent performance of our method in comparison with existing approaches.

2. OUR METHOD

Our method, outlined in Fig. 2, mainly consists of four steps, including DT slicing, local pattern encoding, lacunarity analysis, and feature integration. Each step will be detailed in the following subsections, where bold letters are used for matrices and vectors, regular letters for scalars (such as vector components and dimensions), and calligraphic English alphabets for operators and sets.

2.1. Sequence slicing

A DT sequence is considered as a 3D cube with X, Y, and T axes. Regarding the co-occurrence statistics obtained independently from the XY, XT, and YT planes, we observed certain stationary but distinct spatio-temporal behaviors from the corresponding sampled 2D slices. Thus, we propose to characterize a DT sequence $\boldsymbol{V} \in \mathcal{R}^{M\times N\times K}$ by concatenating appearance-motion information collected from its corresponding 2D slices sampled along three axes. For this purpose, we first apply three slice extracting operators $\mathcal{L}^T$, $\mathcal{L}^Y$, and $\mathcal{L}^X$ to decompose $\boldsymbol{V}$ into slices along the T, Y, and X axes respectively, which are defined as follows,

$$\begin{cases} \mathcal{L}_t^T \circ \boldsymbol{V} = \boldsymbol{V}(:,:,t),\ t = 1,\ldots,K; \\ \mathcal{L}_y^Y \circ \boldsymbol{V} = \boldsymbol{V}(:,y,:),\ y = 1,\ldots,N; \\ \mathcal{L}_x^X \circ \boldsymbol{V} = \boldsymbol{V}(x,:,:),\ t = 1,\ldots,M. \end{cases} \tag{1}$$

With these there operators, three types of 2D slices are obtained, i.e., $\{\mathcal{L}_t^T \circ \boldsymbol{V}\}_{t=1}^K$, $\{\mathcal{L}_y^Y \circ \boldsymbol{V}\}_{y=1}^N$, and $\{\mathcal{L}_x^X \circ \boldsymbol{V}\}_{x=1}^M$.

2.2. Slice-wise pattern encoding

The second step of our method is to locate different kinds of local DT patterns on each 2D slice. Local binary pattern (LBP) is one of the most popular local encoding schemes and has been proved effective for local description in the sense of being invariant to monotonic gray-scale transformations [19]. However the performance of LBP might degrade significantly in practice, especially in the cases of uniform and near-uniform regions, due to its sensitivity to noise. Thus, we use local ternary pattern (LTP) [24] instead. The LTP code on a pixel x, denoted by $\text{LTP}_{P,R}(x)$, is calculated as follow:

$$\text{LTP}_{P,R}(x) = \sum_{p=0}^{P-1} \hat{s}_\tau(g_p - g_x) * 3^p, \tag{2}$$

where the parameters (P, R) define the sampling system with P sampled points within a circular symmetric neighborhood of radius R centered at x, g_x is the gray value of x, g_p is the gray value of the neighbors of x for $(p = 0, 1, ..., P-1)$, τ is a predefined positive constant, $\hat{s}_\tau(y)$ is a threshold function defined as follows,

$$\hat{s}_\tau(y) = \begin{cases} 1, & \text{if } y \geq \tau; \\ 0, & \text{if } |y| < \tau; \\ -1, & \text{if } y \leq -\tau. \end{cases} \tag{3}$$

Though LTP is more resistant to noise than LBP, it generates a huge number of LTP patterns (i.e. 3^P) which leads to high-dimensional features. In this paper, we adopt the modified version of LTP [24] that splits LTP into one positive LBP component and one negative LBP component, which are defined as follows:

$$\begin{cases} \text{LBP}_{P,R}^{+\tau}(x) = \sum_{p=0}^{P-1} s_\tau(g_p - g_x) * 2^p; \\ \text{LBP}_{P,R}^{-\tau}(x) = \sum_{p=0}^{P-1} s_{-\tau}(g_p - g_x) * 2^p, \end{cases} \tag{4}$$

where $s_a(y)$ is a threshold function defined as follows,

$$s_a(y) = \begin{cases} 1, & \text{if } y \geq a; \\ 0, & \text{if } y < a. \end{cases} \tag{5}$$

Given a 2D slice $\boldsymbol{I}$, we define two operators $\mathcal{J}_{P,R}^{+\tau}$ and $\mathcal{J}_{P,R}^{-\tau}$ to generate the code maps $\boldsymbol{C}^+$ and $\boldsymbol{C}^-$ by calculating $\text{LBP}_{P,R}^{+\tau}$ and $\text{LBP}_{P,R}^{-\tau}$ on each voxel of $\boldsymbol{I}$ respectively:

$$\begin{cases} \boldsymbol{C}^+ = \mathcal{J}_{P,R}^{+\tau} \circ \boldsymbol{I}; \\ \boldsymbol{C}^- = \mathcal{J}_{P,R}^{-\tau} \circ \boldsymbol{I}. \end{cases} \tag{6}$$

Benefiting from the employment of the positive and negative LBP encoding schemes, the pattern code maps $\boldsymbol{C}^+$ and $\boldsymbol{C}^-$ corresponding to each 2D slice $\boldsymbol{I}$ enjoy both the robustness to monotonic gray-scale transformations and resistance to random noises and quantization errors.

Given any pattern code map $\boldsymbol{C}$ ($\boldsymbol{C}^+$ or $\boldsymbol{C}^-$) generated from a 2D slice $\boldsymbol{I}$, we partition all its voxels into groups based on the code values and then generate multiple binary images, each of which corresponds to the spatial-temporal distribution of one type of local DT pattern on $\boldsymbol{I}$. In details, we define $\mathcal{M}_p$ as the operator to extract a binary image from a given pattern code map $\boldsymbol{C}$ with respect to the code value p as follows,

$$\mathcal{M}_p \circ \boldsymbol{C}(x) = \begin{cases} 1, & \text{if } \boldsymbol{C}(x) = p; \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Each binary image captures the locations of local DT patterns of certain type in the slice $\boldsymbol{I}$.

2.3. Lacunarity analysis on binarized feature maps

Given the extracted binary images from each 2D slice, the next step is to describe the spatial distribution of each type of local patterns by applying lacunarity analysis on the corresponding binary image. Lacunarity is a powerful tool to measure the spatial distributions of local patterns at a given scale. One widely-used approach to calculating the lacunarity on a binary image $\boldsymbol{B}$ at a certain scale r, denoted as $\Lambda_r(\boldsymbol{B})$, is the gliding box method [25] as follows,

$$\Lambda_r(\boldsymbol{B}) = \frac{E[(X_r^{\boldsymbol{B}})^2]}{(E[X_r^{\boldsymbol{B}}])^2}, \tag{8}$$

where the random variable $X_r^{\boldsymbol{B}}$ denotes the box masses of $\boldsymbol{B}$ at scale r, and $E[\boldsymbol{x}]$ is the expectation value of $\boldsymbol{x}$. We adopt the lacunarity based feature in [26], which states that if the spatial distribution of image patterns of $\boldsymbol{B}$ obeys power law, the corresponding lacunarity should satisfy

$$\ln \Lambda_r(\boldsymbol{B}) = D(\boldsymbol{B}) \ln r + L(\boldsymbol{B}), \tag{9}$$

where $D(\boldsymbol{B})$ and $L(\boldsymbol{B})$ are two scale-independent exponents to encode the scale behavior of lacunarity of $\boldsymbol{B}$ and can be obtained via bi-logarithmic least square fitting with boxes of different sizes. The box sizes are often implemented with a finite sequence of consecutive integers. Then we define an operator $\mathcal{D}$ which applies lacunarity analysis on a binary image $\boldsymbol{B}$ and obtains the corresponding lacunarity-based features as follows:

$$\mathcal{D} \circ \boldsymbol{B} := [D(\boldsymbol{B}), L(\boldsymbol{B})]. \tag{10}$$

The lacunarity-based features are collected from all the binary images as the description for the corresponding 2D slice $\boldsymbol{I}$ by employing the operator $\mathcal{S}_{P,R}^{\tau}$ defined as follows:

$$\begin{aligned} \mathcal{S}_{P,R}^{\tau} \circ \boldsymbol{I} := \Bigg[& \biguplus_{p=0}^{2^P-1} \mathcal{D} \circ \mathcal{M}_p \circ \mathcal{J}_{P,R}^{+\tau} \circ \boldsymbol{I}, \\ & \biguplus_{p=0}^{2^P-1} \mathcal{D} \circ \mathcal{M}_p \circ \mathcal{J}_{P,R}^{-\tau} \circ \boldsymbol{I} \Bigg], \end{aligned} \tag{11}$$

where $\biguplus$ denotes vector concatenation.

2.4. Feature integration over slices

It is observed that the spatial distributions of local DT patterns on 2D DT slices are similar along the same axis. Thus, we calculate the mean vector of the descriptions of all slices along the same axis. This averaging operation both reduces the complexity and enhances the robustness of the resulting features. Finally, our DT descriptor, denoted by $\mathcal{F}$, is defined as the concatenation of all the three mean vectors, i.e.,

$$\begin{aligned} \mathcal{F} \circ \boldsymbol{V} = \Bigg[& \frac{1}{K} \sum_{t=1}^{K} \mathcal{S}_{P,R}^{\tau} \circ \mathcal{L}_t^T \circ \boldsymbol{V}, \\ & \frac{1}{N} \sum_{y=1}^{N} \mathcal{S}_{P,R}^{\tau} \circ \mathcal{L}_y^Y \circ \boldsymbol{V}, \\ & \frac{1}{M} \sum_{x=1}^{M} \mathcal{S}_{P,R}^{\tau} \circ \mathcal{L}_x^X \circ \boldsymbol{V} \Bigg], \end{aligned} \tag{12}$$

where (P, R) and τ are the parameters for both positive and negative LBP coding on all the slices. A common choice of (P, R) and τ are $(8, 3)$ and 5 respectively [24]. Multiple (P, R)s can be used for improvement. For simplicity The parameters are set the same for all slices along different axes.

3. EXPERIMENTAL EVALUATION

In this section, we present a detailed experimental evaluation of our method conducted on two public DT benchmark datasets with several breakdowns.

Table 1. Configurations of different breakdowns of the UCLA and DynTex datasets

Dataset	Breakdown	$\frac{\#\text{Samples}}{\#\text{Classes}}$	#Classes	$\frac{\#\text{Training set}}{\#\text{Classes}}$
UCLA	UCLA-50	4	50	3
	UCLA-SIR	8	50	4
	UCLA-9	4∼108	9	2
	UCLA-8	4∼20	8	2
	UCLA-7	8∼240	7	4
DynTex	Original	10	35	9
	DynTex++	100	36	50
	Alpha	20	3	5
	Beta	7∼20	10	5
	Gamma	7∼38	10	5

3.1. Datasets and configurations

There are mainly two public DT datasets for DT classification: the UCLA dataset [1] and the DynTex dataset [10, 23]. Note that although many static texture datasets have been presented, only a limited number of dynamic texture datasets are available due to the difficulties in collecting DT sequences. To remedy this problem, many studies rearrange the datasets to generate different breakdowns (i.e. sub datasets) for evaluation. The configurations of all these breakdowns used in our experiment are summarized in Table 1. To pursue pure characterization of dynamic texture, the color information of all the experimental data is discarded by converting all color slices of DT sequences to gray-scale images.

Table 2. The classification accuracies (%) of all the compared methods on the test datasets

Method	The UCLA DT Dataset					The DynTex Dataset				
	7-Class	8-Class	9-Class	50-Class	SIR	Basic	DynTex++	Alpha	Beta	Gamma
LBP-TOP [13]	-	-	-	-	-	97.1	89.8	83.3	73.4	72.0
DL-PEGASOS [23]	-	-	95.6	99.0	-	-	63.7	-	-	-
SODM [17]	92.3	-	-	81.0	60.0	-	-	-	-	-
DFS [21]	98.5	99.0	97.5	100	73.8	97.6	89.9	84.9	76.5	74.5
OTF [22]	98.4	99.5	97.2	87.1	67.5	96.7	89.2	-	-	-
WMFS [16]	98.5	97.0	97.1	99.8	61.3	96.5	88.8	-	-	-
Ours	98.1	99.2	96.8	99.7	74.9	97.9	94.8	89.6	80.9	79.9

3.1.1. The UCLA-DT Dataset

The UCLA-DT dataset has been widely used in many previous studies. It originally contains 50 DT categories, each of which contains four gray-scale video sequences captured from different viewpoints. Each video sequence includes 75 frames with 160×110 pixels. For the purpose of reuse as well as adding challenges and reducing ambiguities in evaluating DT classification algorithms, the dataset is reorganized into five different breakdowns, termed UCLA-50 [23], UCLA-SIR [17], UCLA-9 [23], UCLA-8 [7], and UCLA-7 [17]. Interested readers can refer to [21] for more details about the experimental setting on these five breakdowns.

3.1.2. The DynTex Dataset

The DynTex dataset [27] is a diverse collection of high-quality DT videos, which contains more than 650 DT sequences. These video sequences are taken under different environmental conditions involving scaling and rotation, by using static cameras as well as moving ones. The DynTex dataset has been used for evaluating DT classification in many previous studies by different rearrangements, including the original DynTex [13], DynTex++ [23]. Besides, there are three different sub-datasets of the DynTex dataset which have been compiled and labeled for recognition [27], i.e. DynTex Alpha, DynTex Beta and DynTex Gamma. More details can be found in [21, 27].

3.2. Results

We compare our method with six recent DT classification approaches, including LBP-TOP (Local Binary Patterns from Three Orthogonal Planes) [13], DL-PEGASOS (Distance Learning method based on the PEGASOS algorithm) [23], SODM (Spacetime Orientation Distribution Matching) [17], DFS (Dynamic Fractal Spectrum) [21], OTF (Oriented Template based Feature) [22] and WMFS (Wavelet-based Multi-Fractal Spectrum) [16]. The reported results of the compared methods are either available in the literature, or obtained by running the codes which are available online with parameters finely tuned up. The classifier is determined by the protocol on each dataset breakdown. The classification accuracy is reported as the average over a number of trials.

Table 2 summarizes the classification accuracies of all the compared methods on the evaluated datasets. It can be observed from the table that our method is very competitive on the UCLA dataset. Note that with the great development of DT classification techniques, the UCLA dataset is losing its challenging, as it is seen that most of the accuracies of the compared DT descriptors on the UCLA dataset are more than 95%. But at least it demonstrates the comparable performance of our descriptor with other state-of-the-art methods. Among all the breakdowns on UCLA, the SIR test is the most challenging as it requires strong shift robustness on the descriptor. In this case, our method still significantly outperforms all the baseline methods. On the DynTex dataset, our method performs the best. Specifically, our descriptor has significant improvement over the latest DT descriptors on several more challenging datasets (i.e. 4.9% on DynTex++, 4.7% on DynTex Alpha, 4.4% on DynTex Beta, and 4.4% on DynTex Gamma). In summary, the proposed descriptor is very powerful in reliably distinguishing different types of DT sequence.

4. CONCLUSION

We presented a powerful DT descriptor by using lacunarity analysis on positive and negative local binary patterns along two spatial axes and one temporal axis. The proposed method decomposes DT sequence into multiple spatial and temporal binary feature maps and characterizes each feature map with lacunarity analysis. The resulting DT descriptor enjoys both strong robustness and high discrimination. Experiments on several benchmark datasets have demonstrated the power of our method. As can be seen in the experiment, existing datasets for DT classification are very limited. In future, we would like to build up more challenging DT benchmark datasets for DT classification and propose new methodologies for evaluating DT classification algorithms.

5. REFERENCES

[1] G. Doretto, A. Chiuso, Y. Wu, and S. Soatto, "Dynamic textures," *Int. J. of Comput. Vision*, vol. 51, no. 2, pp. 91–109, 2003.

[2] D. Chetverikov and R. Péteri, "A brief survey of dynamic texture description and recognition," in *Comput. Recognition Syst.*, pp. 17–26. Springer, 2005.

[3] P. Saisan, G. Doretto, Y. Wu, and S. Soatto, "Dynamic texture recognition," in *CVPR*. IEEE, 2001, vol. 2, pp. 58–63.

[4] M. Szummer and R. W Picard, "Temporal texture modeling," in *ICIP*. IEEE, 1996, vol. 3, pp. 823–826.

[5] A. B Chan and N. Vasconcelos, "Classifying video with kernel dynamic textures," in *CVPR*. IEEE, 2007, pp. 1–6.

[6] B. Ghanem and N. Ahuja, "Phase based modelling of dynamic textures," in *ICCV*. IEEE, 2007, pp. 1–8.

[7] A. Ravichandran, R. Chaudhry, and R. Vidal, "View-invariant dynamic texture recognition using a bag of dynamical systems," in *CVPR*, 2009, pp. 1651–1657.

[8] G. Doretto, E. Jones, and S. Soatto, "Spatially homogeneous dynamic textures," in *ECCV*, pp. 591–602. Springer, 2004.

[9] F. Woolfe and A. Fitzgibbon, "Shift-invariant dynamic texture recognition," in *ECCV*, pp. 549–562. Springer, 2006.

[10] R. Péteri and D. Chetverikov, "Dynamic texture recognition using normal flow and texture regularity," in *Pattern Recognition and Image Anal.*, pp. 223–230. Springer, 2005.

[11] Z. Lu, W. Xie, J. Pei, and J. Huang, "Dynamic texture recognition by spatio-temporal multiresolution histograms," in *IEEE Workshops on Applicat. of Comput. Vision*. IEEE, 2005, vol. 2, pp. 241–246.

[12] K. G Derpanis, M. Lecce, K. Daniilidis, and R. P Wildes, "Dynamic scene understanding: The role of orientation features in space and time in scene classification," in *CVPR*. IEEE, 2012, pp. 1306–1313.

[13] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, 2007.

[14] G. Zhao, T. Ahonen, J. Matas, and M. Pietikainen, "Rotation-invariant image and video description with local binary pattern features," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1465–1477, 2012.

[15] J. R Smith, C. Lin, and M. Naphade, "Video texture indexing using spatio-temporal wavelets," in *ICIP*. IEEE, 2002, vol. 2, pp. 437–440.

[16] H. Ji, X. Yang, H. Ling, and Y. Xu, "Wavelet domain multifractal analysis for static and dynamic texture classification," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 286–299, 2013.

[17] K. G Derpanis and R. P Wildes, "Dynamic texture recognition based on distributions of spacetime oriented structure," in *CVPR*. IEEE, 2010, pp. 191–198.

[18] K. G Derpanis and R. P Wildes, "Spacetime texture representation and recognition based on a spatiotemporal orientation analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1193–1205, 2012.

[19] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[20] Y. Quan, Y. Xu, and Y. Sun, "A distinct and compact texture descriptor," *Image and Vision Computing*, vol. 32, no. 4, pp. 250–259, 2014.

[21] Y. Xu, Y. Quan, H. Ling, and H. Ji, "Dynamic texture classification using dynamic fractal analysis," in *ICCV*. IEEE, 2011, pp. 1219–1226.

[22] Y. Xu, S. Huang, H. Ji, and C. Fermüller, "Scale-space texture description on sift-like textons," *Comput. Vision and Image Understanding*, vol. 116, no. 9, pp. 999–1013, 2012.

[23] B. Ghanem and N. Ahuja, "Maximum margin distance learning for dynamic texture recognition," in *ECCV*, pp. 223–236. Springer, 2010.

[24] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, 2010.

[25] C. Allain and M Cloitre, "Characterizing the lacunarity of random and deterministic fractal sets," *Phys. Rev. A*, vol. 44, no. 6, pp. 3552–3558, 1991.

[26] Y. Quan, Y. Xu, Y. Sun, and Y. Luo, "Lacunarity analysis on image patterns for texture classification," in *CVPR*. IEEE, 2014, pp. 160–167.

[27] R. Péteri, S. Fazekas, and M. J Huiskes, "Dyntex: A comprehensive database of dynamic textures," *Pattern Recognition Lett.*, vol. 31, no. 12, pp. 1627–1632, 2010.