International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.8, Issue.10, October 2020 E-ISSN: 2347-2693

Dynamic Texture Detection using Flow Estimation based on Texture Constancy

A. Mançour-Billah[1*], E. H. Ait Laasri[2], A. Abenaou[3], D. Agliz[4]

[1,2,4]GEMS Laboratory, National School of Applied Sciences, Ibn Zohr University, Agadir, Morocco
[3]MAISI Laboratory, National School of Applied Sciences, Ibn Zohr University, Agadir, Morocco

[*]*Corresponding Author: anass.mancour@edu.uiz.ac.ma, Tel.: +212600598718*

DOI: https://doi.org/10.26438/ijcse/v8i10.2330 | Available online at: www.ijcseonline.org

Received: 15/Oct/2020, Accepted: 22/Oct/2020, Published: 31/Oct/2020

***Abstract*—** In this work, we aim to detect and classify different dynamic textures representing scenes of outdoor and indoor environments from video sequences. These scenes constitute the vast majority of events in the world, and their detection offers a wide range of applications. Optical flow is one of the most popular methods for motion estimation due to its efficiency and low computational cost. It is based on the brightness constancy assumption, which assumes a constant brightness of the objects between each two frames over time. However, this assumption is not always verified for dynamic textures with non-uniform surface brightness, due to reflections, shadows, transparency or material diffusion. As an alternative, we propose a new flow estimation method based on texture constancy assumption, which describes the spatial texture components motion. The spatial texture of each point of the image, computed using the LBP operator, is assumed to be constant over time. The resulting flow is called texture flow. From its velocity vectors, we extract the magnitude and orientation, which we combine with the texture spatial features to form a shallow hybrid spatiotemporal descriptor. Experimental results on a benchmark database demonstrate both the ability of our method to distinguish between different types of dynamic textures, and its stability with respect to inter and intra-class differences.

***Keywords*—**Dynamic texture, Flow estimation, Motion analysis, Texture constancy, Spatiotemporal descriptors

I. INTRODUCTION

In the world around us, the vast majority of events are in motion (pedestrians, cars in the roads, trees shaken by the wind, water currents in streams and rivers, escalators, etc.). In the image-processing field, these moving events constitute dynamic textures. The latter can be broadly defined as patterns of spatial textures in motion, thus constituting a sequence of images with a certain stationarity in time [1], [2].

A better understanding of the characteristics and properties of these types of textures allows a wide range of applications, such as indoor/outdoor surveillance (fire, flood, traffic, security), autonomous driving, scene understanding and interpretation, motion recognition and segmentation, texture synthesis and content-based video retrieval. These automatic systems, based on visual motion information, are therefore supposed to make more reliable decisions.

In order to describe these dynamic textures, several approaches have been proposed over time. Among them, optical flow is considered as one of the most popular because of its efficiency and low computational cost. Optical flow is defined as the random distribution of the brightness components motion in the image. It serves as an approximation of the real physical movement of the scene by providing a good description of the moving regions [3]. Its estimation is based on the brightness constancy assumption, which supposes a constant brightness of objects between each two frames of the video sequence. However, this assumption is not always verified for dynamic textures whose surface brightness is not always uniform, as it is the case for a wide range of dynamic textures in the environment. These changes in brightness are often due to reflections, shadows, transparency or material diffusion.

In this work, we propose an algorithm for recognition and classification of dynamic textures based on the spatial texture components motion. A new method for motion flow estimation based on the texture constancy assumption is first introduced. In this method, we assume that the spatial texture of each point of the image, computed using the LBP operator, is constant in time. The flow resulting from the motion of these spatial textures is called texture flow. From the vertical and horizontal components of this flow, we compute the magnitude and orientation of the texture motion between each two frames. Two features representing the two spatial and temporal modes are then extracted: (1) for the spatial mode, we use the LBP histogram by cumulating the magnitude of the texture flow; (2) for the temporal mode, we use the histogram of the orientations of the proposed flow by once again cumulating the magnitude. For a local representation, these histograms are computed from spatiotemporal windows of defined sizes called cuboids, then each cuboid is individually classified by feeding it to a previously trained SVM classifier. For a first experiment, we apply our approach to the DynTex database [4], where we extract the confusion matrix from a number of dynamic texture classes. We then compare our method to other baseline methods and their hybridization.

The rest of the paper is organized as follows. Section 2 gives an overview of previous work related to dynamic texture detection. Section 3 provides a brief description of used methods and the proposed approach. Experimental results for the evaluation of our algorithm are then presented in section 4. Finally, section 5 concludes the article.

II. LITERATURE REVIEW

In recent years, research on dynamic textures has aroused the interest of the computer vision community mainly because of the many potential applications. Several approaches have been proposed over time, and can be briefly classified into model-based, discrimination-based and motion-based methods.

Model-based methods aim at building models founded on generative processes for dynamic texture description [1], [2], [5]. The parameters extracted from these dynamic textures are then classified according to the models that represent them. Doretto et al. [6] used the Gauss-Markov model for parameter modeling and estimation. This model is based on a linear dynamic system (LDS) which explores the spatial and temporal regularities of dynamic textures. Chan and Vasconcelos [7] proposed the expectation-maximization algorithm to extract the parameters of dynamic textures and classify them. The authors then represented the video sequences using a layered dynamic texture as a collection of stochastic layers. Although these methods are considered quite robust, their applications to local-scale moving scenes proved to be less efficient.

On the other hand, discrimination-based methods rely on the statistical properties of the spatial distribution of pixels in the video sequence. Among the most widely used approaches are those based on Local Binary Patterns (LBP) [8], which were initially proposed for static textures and later extended to dynamic textures. The strength of this method lies in its ability to describe the texture locally by applying the comparison of each pixel with its neighborhood, independently of the compared values. Extensions of the LBP include Uniform LBP [9], Completed LBP [10], [11] and Extended LBP [12], [13]. Among its spatiotemporal extensions are the Volume Local Binary Patterns (VLBP) [14], which consider the entire three-dimensional neighborhood of the pixel; and the Local Binary Patterns from Three Orthogonal Plans (LBP-TOP) [15], which compute the LBP features in the three orthogonal planes XY, XT and YT. Nevertheless, these methods have the disadvantage of having quite large histogram sizes, ranging from 2^{14} to 2^{26}, which is not very suitable for real-time applications.

Finally, motion-based methods are the most popular because of their efficiency and low computational cost. These methods generally extract the motion properties of the dynamic texture from the optical flow. In [16], the authors have mapped the amplitudes and directions of the normal flow as spatiotemporal textures, allowing a representation where the spatial and temporal aspects of the texture are coupled. Péteri and Chetverikov [17] used the normal flow and the texture regularity to extract quantitative features such as orientation, divergence and periodicity. Fazekas and Chetverikov [18] adopted an analysis of local image distortions from optical flow for the extraction of scale-invariant and rotation-invariant features. Fazekas et al. [19] also addressed the so-called strong dynamic textures (having intrinsic dynamics) by evaluating three alternative methods of optical flow estimation, namely: gradient constancy, color constancy and brightness conservation (when an object can diffuse its brightness to its neighborhood). The brightness conservation proved to be the most adequate. Optical flow has also been combined with other discriminative methods in Chen et al. [20]. LBP and WLD (Weber Local Descriptor) were used for both spatial and temporal modes, and were combined with the Histogram of Oriented Optical Flow (HOOF) calculated from the optical flow. HOOF, being an optical flow-based approach describing the movement properties at each moment, has been widely used in events and human actions recognition [21], [22], [23]. Recently, Kaltsa et al. [24] have proposed an algorithm for recognizing and localizing dynamic textures in outdoor environments. The presented descriptor, called LBP flow, combines the features of the LBP with those of the optical flow by calculating the LBP of the optical flows obtained from the three orthogonal planes XY, XT and YT. However, motion-based methods rely on two assumptions that are not always verified in dynamic textures of a stochastic nature: local regularity, where the projected motion is assumed to be parallel to the image plane; and brightness constancy, which assumes a constant brightness of the objects between each two frames through time. In the case where the brightness is non-uniform, especially in dynamic textures containing a wide diversity of luminosities and colors, approaches based on optical flow reach their limits.

III. PROPOSED METHOD

The proposed method is divided into two main steps: the estimation of the flow and then the extraction of its features. For the flow estimation, we present a new technique based on the texture constancy assumption, where the texture between two frames is assumed to be constant in time. From the estimated flow, we extract two features, a spatial one based on the local texture and a temporal one based on its orientations.

A. Texture Flow Estimation

Optical flow is the random distribution of the brightness components motion constituting a given image. Its application to a video sequence results in a field of velocity vectors that gives a general idea of the direction and magnitude of the motion. Thus, the optical flow serves as an approximation of the real physical movement of the scene and allows a good description of the moving regions [3].

The optical flow principle is based on the brightness constancy assumption, i.e. the brightness of each point of the image is assumed to be constant over time. Thus, between each two frames of the video sequence, the brightness of a pixel moves according to two horizontal and vertical components according to the following equation:

$$I_x u + I_y v + I_t = 0 \qquad (1)$$

With I_x and I_y are the spatial derivatives of the image brightness I, and I_t its temporal derivative. u and v are the two horizontal and vertical components of the optical flow (velocity vectors).

In order to solve this equation with two unknowns, different methods have been proposed. Among these is the Lucas-Kanade method [25], which has been used successfully in recent years. The method assumes that the displacement of the brightness components between two consecutive frames is small, and is approximately constant in a spatial neighborhood of a considered point. It is therefore possible to assume that the optical flow equation is valid for all the pixels of a given window, which makes it possible to extract the velocity vectors. These velocity vectors are then used to calculate the orientation and magnitude of the motion.

However, the optical flow is based on the brightness constancy assumption, which means that the brightness of an object is assumed to remain unchanged from frame to frame. This condition is not always satisfied in natural dynamic textures, as in surfaces with non-uniform brightness because of reflections, shadows, transparency or material diffusion. An example of this type of dynamic textures is water, whose appearance changes with location and external conditions. Figure 1 shows different aspects of the same dynamic texture with a non-uniform surface brightness.

Figure 1. Different aspects of water as a dynamic texture with a non-uniform surface brightness (DynTex database [4]).

Other alternatives to brightness constancy have been proposed in the literature, among which are gradient constancy [26] and color constancy [27].

Gradient constancy was used to bypass changes in brightness in the image. The optical flow constraint equation based on the gradient constancy assumption (using Laplacian here) is given as follows:

$$I_{xxt} + I_{yyt} + u(I_{xxx} + I_{yyx}) + v(I_{xxy} + I_{yyy}) = 0 \qquad (2)$$

Color consistency has also been used to overcome problems caused by brightness changes. The optical flow constraint equations based on color constancy are given as follows:

$$\begin{aligned} I_t^r + uI_x^r + vI_y^r &= 0 \\ I_t^\phi + uI_x^\phi + vI_y^\phi &= 0 \\ I_t^\theta + uI_x^\theta + vI_y^\theta &= 0 \end{aligned} \qquad (3)$$

With r, φ and θ the components of spherical coordinates in RGB space.

We propose in this work a new method for flow estimation based on texture. The estimation of the flow will be based only on the texture constancy assumption.

As with brightness and color, texture is considered as an important feature for object recognition and description in the image. In the case of natural dynamic textures, we notice that they can change both brightness and color, but generally keep their basic textures. Figure 2.a shows an example of a dynamic texture with non-uniform surface brightness. This texture contains two regions of different illuminations. Based on the brightness constancy assumption for optical flow estimation, the two regions would have velocity vectors of different magnitudes, the larger of which would correspond to the regions of high brightness. After the application of the LBP in figure 2.b, only the texture is taken into account, and we notice a homogenization of the aspect of the dynamic texture surface, independently of its brightness.

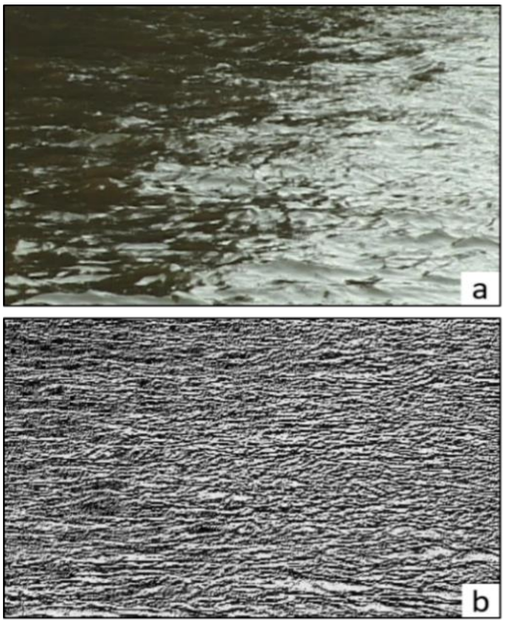

Figure 2. (a) Frame of a dynamic texture with a non-uniform surface brightness, and (b) its LBP image.

Based on this principle, we believe that the texture flow, calculated from surfaces with different illuminations or colors but belonging to the same dynamic texture, will be described with approximately similar velocity vectors.
First, we apply the LBP operator to each frame of the video sequence to convert the brightness image into a textured image. The LBP is a discriminative method that acts locally by comparing each pixel value with its neighborhood. This comparison provides a binary sequence that describes the texture of each pixel independently of the rate of the compared values, thus providing invariance to brightness.

The LBP equation is given as follows:

$$LBP_{P,R} = \sum_{p=0}^{p-1} s\left(g_p - g_c\right)^{2p} \tag{4}$$

Where $s(x)$ is the sign function, and g_c and g_p are the values of the pixel to be evaluated and those of its neighborhood respectively. The binary sequence is multiplied by the binomial weight of each neighbor (2^p), to obtain the LBP code of the pixel.

Under the texture constancy assumption, we assume that the LBP code of each pixel is constant in time and that it moves spatially along two components u and v. This assumption can be formulated as follows:

$$LBP_I(x+u, y+v, t+1) = LBP_I(x, y, t) \tag{5}$$

Where LBP_I is the LBP image calculated from the image brightness I.

Applying the first-order Taylor approximation of the above equation, we obtain the following flow equation:

$$LBP_{I_t} + u.LBP_{I_x} + v.LBP_{I_y} = 0 \tag{6}$$

Where LBP_{I_t} is the temporal derivative of the LBP image, whereas LBP_{I_x} and LBP_{I_y} are its horizontal and vertical spatial derivatives respectively.

In order to solve this equation with two unknowns, we apply the same Lucas-Kanade method previously presented for the brightness constancy to extract the velocity vectors.

The magnitude and orientation of the flow are then calculated as follows:

$$m = \sqrt{u^2 + v^2} \tag{7}$$

$$\theta = \tan^{-1}\left(\frac{v}{u}\right) \tag{8}$$

Figure 3 shows the LBP images of two consecutive frames of a video sequence and their resulting flow. The video sequence is obtained from the Video Water database [28]. The orientations of the arrows show the orientation of the flow and their sizes show its magnitude. We notice that the resulting flow is characterized by homogeneity over the entire dynamic surface, and that it was able to uniformly describe the movement of water on both bright regions and those containing reflections.

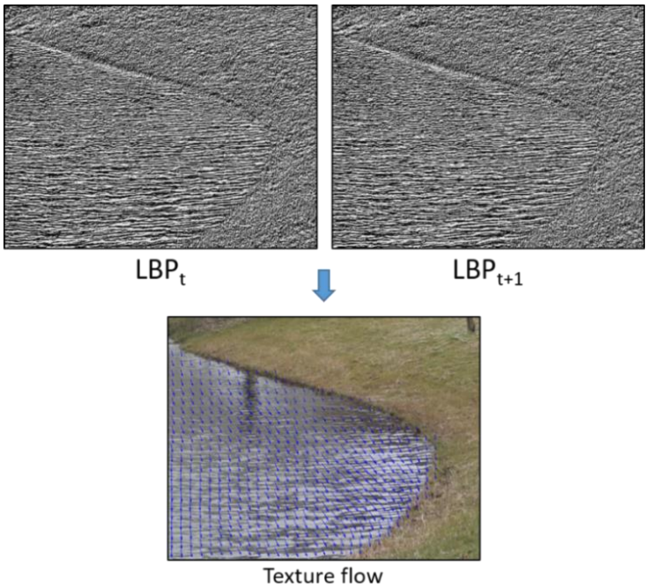

Figure 3. LBP images of two consecutive frames at time t and their resulting flow (Video Water database [28]).

B. *Extraction of Flow Features*

Dynamic texture can be recognized by both its dynamics and its physical appearance. The texture flow previously introduced has the particularity to describe the physical aspect of the dynamic texture as well as the dynamics of this aspect. First, the video sequence is divided into n × m × t non-overlapped regions. For each of these cuboids, we extract the features of the two spatial and temporal modes.

1) *Spatial Features*

For the spatial mode, we extract the LBP histogram at each time t for each n × m block of the frame. This histogram contains 2^P bins, where P is the number of neighbors used for the calculation of the LBP. Each pixel of the LBP image block contributes with its magnitude m (7) obtained through the texture flow. The equation of the histogram extracted for each instant t is given as follows:

$$H(k) = \sum_{i=1}^{n} \sum_{j=1}^{m} f(LBP_I(i,j), k),\ k \in [0, 2^P]$$
$$with \quad f(x,y) = \begin{cases} m(i,j) & \text{x = y} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

These histograms are then concatenated to obtain a spatial descriptor vector of size $2^P \times t$.

2) Temporal Features

Inspired by the success of the Histograms of Oriented Optical Flow (HOOF) [21], which is an optical flow-based approach describing the motion characteristics at each instant, we use a similar descriptor based on texture flow, which we call Histogram of Oriented Texture Flow (HOTF).

For each instant t in an n × m block, the magnitudes of the texture flow are cumulated in a histogram of orientations θ (8) using B bins. The orientation of each pixel of the block is assigned in a bin as follows:

$$-\frac{\pi}{2} + \pi\frac{b-1}{B} \leq \theta(i,j) < -\frac{\pi}{2} + \pi\frac{b}{B},\ b \in [1, B] \tag{10}$$

The histograms obtained are then concatenated to obtain a temporal descriptor vector of size B × t.

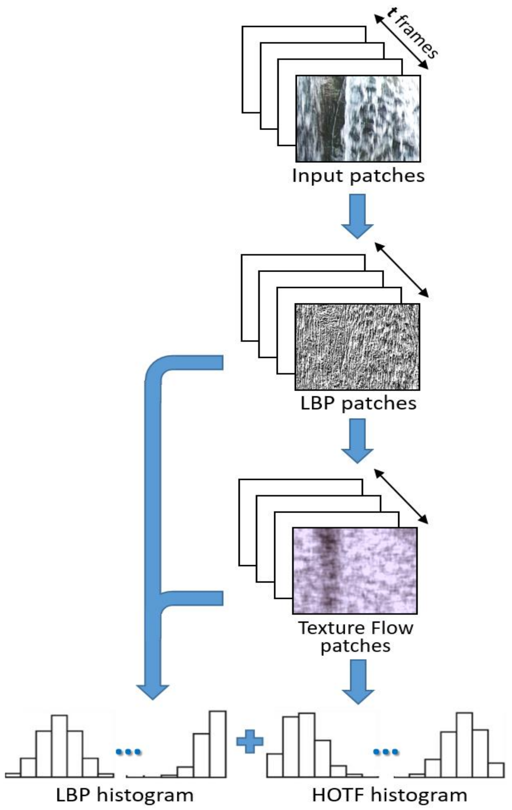

Figure 4. Feature extraction scheme from the texture flow applied to an n × m sized patch during t consecutive frames.

Finally, the two spatial and temporal descriptor vectors are concatenated and then normalized to form a shallow spatiotemporal descriptor vector of size $(2^P + B) \times t$ for each cuboid n × m × t.

Figure 4 shows the feature extraction scheme from the texture flow applied to a patch of size n × m during t consecutive frames.

IV. EXPERIMENTAL RESULTS

We first start by defining the parameters that we will use during the experiment in which we evaluate our method on the DynTex benchmark database [4]. For this, we use the different dynamic texture classes of the database, where each class is subdivided into two thirds for training and the other third for testing.

A. Parameters Settings

To estimate the texture flow, two parameters are involved: the number of neighbors P, used for the calculation of the LBP operator of each frame; and the size of the neighborhood of each point, which is assumed to be constant during the flow estimation using the Lucas-Kanade method. In order to assimilate the texture flow to the standard optical flow which is based on brightness varying from 0 to 255, we have defined the number of neighbors P to 8 immediate neighbors, thus allowing the LBP codes to also vary between 0 and 255. For the flow estimation window, the use of a small window has the ability to capture the most subtle movements, but risks missing larger movements. Contrary, a larger window shows an inverse behavior, which gives some resistance to occlusions, but increases considerably the computational cost. As a compromise between these two behaviors, a size of 21 × 21 was adopted when estimating the texture flow.

For the extraction of spatial and temporal features, the cuboid size n × m × t constitutes a compromise among the descriptors accuracy, their local or global nature, and the computational cost. A small spatial window will be able to describe the dynamic texture in a more local way, but may not contain enough details for a proper discrimination; while a larger spatial window may confuse different textures of the same scene or ignore some of them. Since the emphasis here is on recognition rather than segmentation, we choose for the experiment a spatial window with dimensions n = m = 101. The time window also plays an important role in the accuracy of the descriptors. A high number of considered frames results in larger descriptor vectors, containing more information for better discrimination, but increasing the computational cost. For each of the videos in the database, we set t = 10 frames only.

For the extraction of the flow orientations histogram, a high number of bins describes the most subtle dynamic texture movements, but this is likely to generate noise; while a lower number describes the movement in a coarser way. We use in this work 24 orientations spaced by 15°.

As output, we obtain a spatiotemporal descriptor vector of size $(2^8 + 24) \times 10$ for each patch of size 101×101. This vector is then passed to an SVM classifier for individual classification.

B. Database

For the evaluation of our algorithm, we used the renowned DynTex database [4]. It is one of the main databases dealing with dynamic texture classification. It contains a wide variety of high quality videos divided into several classes, sometimes with high intra-class variance. These classes represent dynamic textures found mainly in outdoor environments (sea, trees, automobile traffic, flags, etc.), but also indoor environments (escalators). Each class is divided into two thirds for training and one third for testing.

Figure 1 shows examples of dynamic textures extracted from DynTex database. Although these textures belong to different classes (sea, calm water, fountains), they can also be classified under a single mother class (water).

C. Recognition

The recognition experiments give a general idea of the algorithm's performance in extracting the different characteristics constituting the spatiotemporal behavior of each dynamic texture, as well as its ability to distinguish them. For this purpose, we evaluate the accuracy of the algorithm by assigning the video sequences from the test data to their corresponding representational categories. Each input video is thus associated to a unique class according to its SVM output.

Eight classes from DynTex database were used for the experiment, namely: flag, grass, trees, fountains, sea, calm water, escalator, and traffic. These classes essentially represent dynamic textures found in outdoor environments, and can also be subdivided into subclasses belonging to more general classes.

Table 1 presents the confusion matrix of the classification accuracies obtained for these eight classes. As indicated, our method achieves high accuracy for 6 out of the 8 classes. The classes containing flags, fountains, escalators and traffic are perfectly distinguished and are not confused with any other dynamic texture, although they may have roughly similar temporal behaviors.

Table 1. Confusion matrix of the classification accuracies of eight classes from DynTex database.

	Flag	Grass	Trees	Fountains	Sea	Calm water	Escalator	Traffic
Flag	100%							
Grass		100%						
Trees		14.3%	85.7%					
Fountains				100%				
Sea					100%			
Calm water					14.3%	85.7%		
Escalator							100%	
Traffic								100%

This is explained by the nature of our method which focuses on the motion of the local spatial behavior of the texture, rather than its apparent motion. Moreover, it combines these spatial and temporal behaviors to describe the dynamic texture in a spatiotemporal way. Thus, dynamic textures with similar dynamics would not be confused since they would look different, and vice-versa.

However, this allows our algorithm to overcome intra-class differences that may be encountered within some more general classes. As mentioned in the table, misclassifications of some classes refer to other classes of the same nature. The two classes 'calm water' and 'sea', which have been confused, can in fact be considered sub-classes of the more general 'water' class, which indicates that our algorithm has been able to recognize the spatiotemporal behavior of water. Although fountains also represent water, they have not been confused because their spatiotemporal behavior is usually different from that of surface water (spouting, dripping). On the same principle, the 'trees' class has also been confused with the 'grass' one, both of which may belong to the general class 'vegetation'.

We then compare our method to other similar methods also used. Table 2 includes two baseline methods and their hybridization, namely: the Histogram of Oriented Optical Flow (HOOF) [21] and the LBP operator [8]. For the feature extraction of each of these methods, we used spatiotemporal windows of the same size, and an identical number of bins equal to 24 for both HOOF and HOTF (Histogram of Oriented Texture Flow).

It is easily noted that the results are less efficient when using a descriptor acting on a single dimension, in particular the LBP for the spatial dimension and the HOOF for the temporal one. However, during their hybridization, dynamic textures are more adequately distinguished. Indeed, a dynamic texture class can usually resemble another either spatially or temporally, but hardly spatiotemporally. Nevertheless, LBP+HOOF is characterized by a weak intra-class distinction. As shown in the table, the two classes 'sea' and 'calm water', characterized by a high similarity, are generally misclassified, both with an accuracy of 71.43%, against 100% and 85.71% respectively when hybridizing with the texture flow (LBP+HOTF). This difference between the two hybridizations is mainly due to the time component. Different from the optical flow which describes the motion of the apparent brightness, the texture flow describes the motion of the spatial texture itself, which leads to a temporal descriptor taking into account the most subtle changes of the texture over time, regardless of changes in illumination or color. Thus, this property allows a more precise distinction between classes of similar apparent behavior.

As an overall average accuracy, the proposed method achieves a score of 95.83% on the totality of videos used for the test, outperforming the baseline methods. This proves the robustness of our approach as well as its stability with respect to inter and intra-class differences. In addition, the use of only 10 frames when extracting temporal features has little effect on accuracy, but significantly reduces the size of the descriptor, making the approach suitable for real-time applications.

Table 2. Comparison with baseline methods and their hybridization of eight classes from DynTex database.

DT classes	HOOF	LBP	LBP+HOOF	LBP+HOTF
Sea	57.14	42.86	71.43	100
Grass	71.43	57.14	100	100
Trees	57.14	85.71	85.71	85.71
Flags	57.14	100	100	100
Calm water	85.71	71.43	71.43	85.71
Fountains	85.71	71.43	100	100
Escalator	100	50	100	100
Traffic	75	100	100	100
Average	70.83	72.92	89.58	**95.83**

V. CONCLUSION

In this work, we addressed the problem of recognition and classification of dynamic texture based on the spatial texture components motion. We first introduced a new technique for estimating the motion flow based on the texture constancy assumption. This assumption consists in assuming that the spatial texture of each pixel of the frame is constant in time. The texture was computed using the LBP operator, then its flow was estimated using the Lucas-Kanade method. The resulting flow was called texture flow. From the velocity vectors of this flow, we extracted the magnitude and orientation, which we combined with texture spatial features to form a shallow hybrid spatiotemporal descriptor. For each spatiotemporal window, we cumulated the magnitudes of the texture flow in two histograms: the LBP histogram for the spatial mode, and the Histogram of Oriented Texture Flow (HOTF) for the temporal mode. Each resulting vector was fed to an SVM classifier individually.

We then tested our method on the DynTex benchmark database using eight classes. The obtained accuracies reached high scores for the majority of the classes, with an overall score of 95.83%. We also compared our method to other baseline methods such as LBP, HOOF and their hybridization, by applying them on the same classes and using the same parameters. It seems clear that our method allows a better classification due to its ability to overcome intra-class variations, which makes it stable towards both inter and intra-class differences. Moreover, the reduced size of the spatiotemporal descriptor makes our approach suitable for real-time applications.

REFERENCES

[1] D. Chetverikov and R. Péteri, "*A brief survey of dynamic texture description and recognition*," In: Computer Recognition Systems. Springer, p. **17–26**, **2005**.

[2]G. Doretto, E. Jones and S. Soatto, "*Spatially homogeneous dynamic textures*," In: Computer Vision-ECCV 2004. Springer, pp. **591-602**, **2004**.

[3] P. Turaga, R. Chellappa and A. Veeraraghavan, "*Advances in video-based human activity analysis: challenges and approaches*," Advances in Computers, vol. **80**, pp. **237-290**, **2010**.

[4] R. Péteri, S. Fazekas and Mark J. Huiskes, "*DynTex: A comprehensive database of dynamic textures*," Pattern Recognition Letters, vol. **31**, Issue. **12**, pp. **1627-1632**, **2010**.

[5]F. Yang, G.-S. Xia, G. Liu, L. Zhang and X. Huang, "*Dynamic texture recognition by aggregating spatial and temporal features via ensemble {SVMs}*," Neurocomputing, vol. **173**, pp. **1310-1321**, **2016**.

[6]G. Doretto, A. Chiuso, Y. Wu and S. Soatto, "*Dynamic Textures*," International Journal of Computer Vision, vol. **51**, Issue. **2**, pp. **91-109**, **2003**.

[7] A. Chan and N. Vasconcelos, "*Modeling, Clustering, and Segmenting Video with Mixtures of Dynamic Textures*," IEEE transactions on pattern analysis and machine intelligence, vol. **30**, Issue. **5**, pp. **909-926**, **2008**.

[8] T. Ojala, M. Pietikainen and T. Maenpaa, "*Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*," Pattern Anal. Mach. Intell., vol. **24**, Issue. **7**, p. **971–987**, **2002**.

[9] O. Lahdenoja, J. Poikonen and M. Laiho, "*Towards understanding the formation of uniform local binary patterns*," ISRN Machine Vision, **2013**.

[10] Z. Guo, L. Zhang and D. Zhang, "*A completed modeling of local binary pattern operator for texture classification*," IEEE Transactions on Image Processing, vol. **19**, Issue. **6**, p. **1657–1663**, **2010**.

[11] D. Tiwari and V. Tyagi, "*Dynamic texture recognition based on completed volume local binary pattern*," Multidimensional Systems and Signal Processing, vol. **27**, pp. **563-575**, **2016**.

[12]L. Liu, L. Zhao, Y. Long, G. Kuang and P. Fieguth, "*Extended local binary patterns for texture classification*," Image and Vision Computing, vol. **30**, p. **86–99**, **2012**.

[13] R. Maurya, "*Melanoma Detection Using Modified Extended LBP*," International Journal of Computer Sciences and Engineering, vol. **6**, Issue. **7**, pp. **698-703**, **2018**.

[14]G. Zhao and M. Pietikäinen, "*Dynamic texture recognition using volume local binary patterns*," In: Dynamic Vision. Springer, pp. **165-177**, **2007**.

[15]G. Zhao and M. Pietikäinen, "*Dynamic texture recognition using local binary patterns with an application to facial expressions,*" IEEE transactions on pattern analysis and machine intelligence, vol. **29**, Issue. **6**, pp. **915-928**, **2007**.

[16] C.-H. Peh and L.-F. Cheong, "*Synergizing spatial and temporal texture,*" IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, pp. **1179-1191**, **2002**.

[17]R. Péteri and D. Chetverikov, "*Dynamic Texture Recognition Using Normal Flow and Texture Regularity,*" Iberian Conference on Pattern Recognition and Image Analysis, pp. **223-230**, **2005**.

[18] S. Fazekas and D. Chetverikov, "*Dynamic texture recognition using optical flow features and temporal periodicity,*" In: 2007 International Workshop on Content-Based Multimedia Indexing, pp. **25-32**, **2007**.

[19] S. Fazekas, T. Amiaz, D. Chetverikov and N. Kiryati, "*Dynamic Texture Detection Based on Motion Analysis,*" International Journal of Computer Vision, vol. **82**, pp. **48-63**, **2009**.

[20]J. Chen, G. Zhao, M. Salo, E. Rahtu and M. Pietikainen, "*Automatic dynamic texture segmentation using local descriptors and optical flow,*" IEEE transactions on image processing: a publication of the IEEE Signal Processing Society, vol. **22**, Issue. **1**, pp. **326-339**, **2013**.

[21] R. Chaudhry, A. Ravichandran, G. Hager and R. Vidal, "*Histograms of oriented optical flow and Binet-Cauchy kernels on nonlinear dynamical systems for the recognition of human actions,*" In: IEEE Conference on Computer Vision and Pattern Recognition, pp. **1932-1939**, **2009**.

[22]J. Pers, V. Kenk, M. Kristan, M. Perše, K. Polanec and S. Kovačič, "*Histograms of optical flow for efficient representation of body motion,*" Pattern Recognition Letters, vol. **31**, pp. **1369-1376**, **2010**.

[23] S. S and A. E. Mendonca, "*Unusual Events Detection via Global Optical Flow and SVM,*" International Journal of Computer Sciences and Engineering, vol. **4**, Issue. **3**, pp. **179-183**, **2016**.

[24]V. Kaltsa, K. Avgerinakis, A. Briassouli, I. Kompatsiaris and M. Strintzis, "*Dynamic texture recognition and localization in machine vision for outdoor environments,*" Computers in Industry, vol. **98**, **2018**.

[25] B. Lucas and T. Kanade, "*An iterative image registration technique with an application to stereo vision (IJCAI),*" Proceedings of Imaging Understanding Workshop, pp. **121-130**, **1981**.

[26] N. Papenberg, A. Bruhn, T. Brox, S. Didas and J. Weickert, "*Highly Accurate Optic Flow Computation with Theoretically Justified Warping,*" International Journal of Computer Vision, vol. **67**, pp. **141-158**, **2006**.

[27] Y. Mileva, A. Bruhn and J. Weickert, "*Illumination-Robust Variational Optical Flow with Photometric Invariants,*" Pattern Recognition, pp. **152-162**, **2007**.

[28]P. Mettes, R. T. Tan and R. C. Veltkamp, "*Water detection through spatio-temporal invariant descriptors,*" Computer Vision and Image Understanding, vol. **154**, pp. **182-191**, **2017**.

AUTHORS PROFILE

Anass Mançour-Billah received his Bachelor of Engineering degree in Computer Science from the National School of Applied Sciences, Ibn Zohr University, Agadir, Morocco, in 2014. He is currently pursuing Ph.D. in Image and Signal Processing at Ibn Zohr University, Morocco. His research interests comprise image processing, computer vision, machine learning and optimization.

El Hassan Ait Laasri is an assistant professor of Electronics at Ibn Zohr University, Agadir, Morocco. His research interests lie in the field of instrumentation and signal & image processing using both classical processing methods and artificial intelligence techniques. During his thesis, he had worked on two monitoring systems: A structural health monitoring system (Italy) and a seismic monitoring system (Morocco). Several tasks were developed based on a variety of signal processing methods and artificial intelligence techniques such as neuronal network and fuzzy logic.

Abdenbi Abenaou was born in Casablanca, Morocco. He is graduated from The Bonch-Bruevich St. Petersburg State University of Telecommunications, Saint Petersburg, Russia. He obtained his Ph.D. degree in Technical Sciences from the St. Petersburg State Electrotechnical University "LETI", in 2005. He is currently a professor at the department of Software Engineering in National School of Applied Sciences, Agadir, Morocco. His research interests lie in the field of signal and image processing.

Driss Agliz was born and brought up in Morocco. He studied Physics and Chemistry, and holds, in 1983, his BSc degree (LicenceEs-Sciences Physique) in Physics from the University Cadi Ayyad, Marrakech, Morocco. He then joined Rennes I University – France, where he got, in 1984, a MSc. degree (DEA: Diplôme des Etudes Approfondies) in materials sciences. He pursued his research on the studies and got, in 1987, a PhD degree from the University of Rennes I - France. Dr. D. Agliz started his teaching career on 1987 as Assistant Professor in the faculty of Science at University Ibn Zohr. Then he holds a "Doctorat d'Etat" thesis in 1997. He is currently Full Professor in National School of Applied Sciences, University Ibn Zohr, Agadir, Morocco. He has been working in several fields, such signal processing, materials sciences. He is supervising PhD, MSc as well as BSc students in the field of signal processing and renewable energies.

