DynTex: A Comprehensive Database of Dynamic Textures ⋆

Renaud Péteri [a,*], Mark J. Huiskes [b] Sándor Fazekas [c]

[a] *Mathematics, Image and Applications Laboratory (MIA)*
University of La Rochelle, avenue Michel Crépeau, 17000 La Rochelle, France

[b] *Leiden Institute for Advanced Computer Science (LIACS)*
Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

[c] *MTA SZTAKI - Hungarian Academy of Sciences*
1111 Budapest, Kende u. 13-17., Hungary

Abstract

We present the DynTex database of high quality dynamic texture videos. It consists of over 650 sequences of dynamic textures, mostly in everyday surroundings. By analyzing the nature and structure of dynamic textures, we were able to annotate the videos by their semantic content and the visual properties of the physical processes underlying the dynamical textures. The videos and annotations are made publicly available for scientific research.

Key words: Dynamic texture, scientific image collections, video segmentation, video processing, pattern recognition, annotation.
PACS: 97.10.Gz(change!!!), 97.30.Qt, 97.80.Gm

1 Introduction

In recent years we have witnessed a rapid growth of interest in the study of *dynamic texture* (DT). This new field of research offers an extension of the study of static texture into the temporal domain. Dynamic texture phenomena can be observed all around us in daily life, e.g. moving trees in the wind,

⋆ http://www.cwi.nl/projects/dyntex/

* Corresponding author.
Email addresses: renaud.peteri@univ-lr.fr (Renaud Péteri), markh@liacs.nl (Mark J. Huiskes), sandor.fazekas@sztaki.hu (Sándor Fazekas).

rippling water, fluttering sails and flags, or moving crowds of people through a shopping street.

Just like static textures, dynamic textures can be studied in a wide variety of ways. People working on *texture synthesis* try to model dynamic textures for realistic rendering (e.g. [13], [11] or [7]). A related problem is *texture analysis* where the problem is to characterize the visual properties of the texture, e.g. its regularity. This can also be useful for texture retrieval. For *texture detection* the goal is to detect when a certain type of dynamic texture appears in a video sequence, e.g. to detect fire ([5]). An interesting problem is also to discern dynamic texture from camera motion ([1]). *Texture segmentation* is about video segmentation and the accurate localization of the textures, in space, and, possibly, in time (e.g. [6,2]). For *texture recognition* the aim is to recognize the type of dynamic texture, possibly from among several other ones [10,14]; A review on DT description and recognition is presented in [3]. In most of the cases, works on this topic consider well segmented sequences of dynamic textures. A final topic, *irregularity detection*, that to our knowledge has not been investigated so far, aims at detecting irregular motions in sequences containing pure dynamic textures (for instance the detection of a leaf drifting on a river surface). It can be seen as the counterpart of defect detection for static textures (see [4]).

With respect to datasets, each of these areas of study comes with its own requirements for performance evaluation. For instance, analysis for texture synthesis may best be done on close-up sequences of the texture, whereas detection requires that the dynamic texture is shown in its context. For texture segmentation it may be beneficial to offer sequences where several dynamic texture phenomena are present in the same sequence.

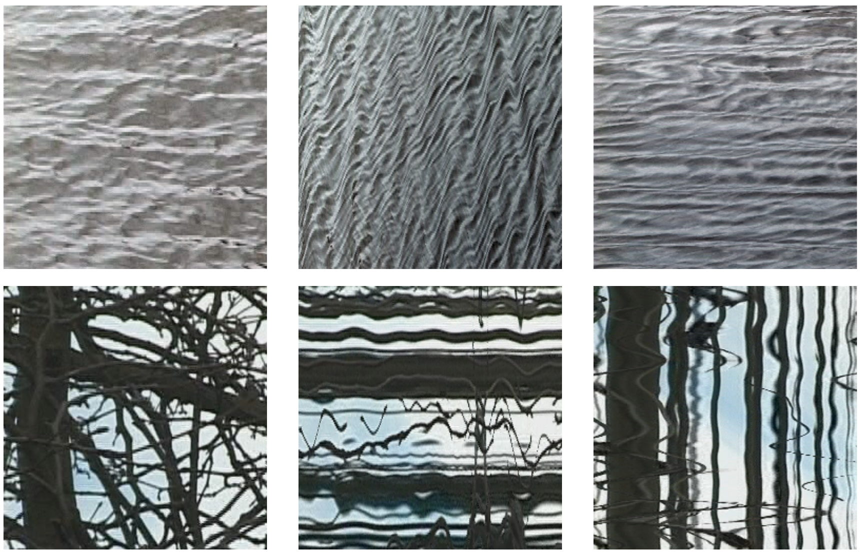
In this paper we present the DynTex database. It consists of over 650 high-quality sequences of dynamic texture. It aims to serve as a standard database for dynamic texture research and to accommodate the needs for assessing the different research issues mentioned above. So far no other databases suitable for this purpose are available. One interesting and pioneer database to mention is the dataset of the MIT [12]. It is composed of around 25 black and white segmented sequences of dimension 170x115x120. However, this collection has a number of drawbacks: videos dimensions are small (especially in the temporal direction); there is only a single occurrence per class, and not enough classes are available for practical classification purposes; finally, some of the sequences show undesirable camera motion. Other datasets of dynamic textures used in research papers have been shot by the authors themselves and do not aim at being publicly released, which prevents from using them for comparison with other research works.

The need for a standard database is clearly demonstrated by the interest of the research community: at the time of writing of this paper, DynTex has already over 150 registered researchers (about half of which are PhD candidates) that use the database for their studies.

Mark? [Structure of the paper. In Section 2 we discuss [not clear yet;possibly the basic structure of the database, or design philosophy] In Section 1 we discuss the acquisition protocal of the texture sequences and describe the video formats in which the sequences are offered. In Section 4 we describe the annotation scheme of the sequences. In Section 5 we discuss how users of the database can create their own test sets from the total DynTex set.]

2 Dynamic Textures

Giving a proper definition about what is a texture is very difficult. The reason is that texture can have many aspects, from regular to stochastic, they can be orientated or isotropic, and can occur at a given scale or have a fractal behavior. The French scientist Yves Meyer once described texture as a subtle balance between repetition and innovation [1] .



(a) (b) (c)

Fig. 1. The 'river' sequence (1^{st} row) and the 'tree' sequence (2^{nd} row). (a) First frame of the sequence. (b) YT cut at $X = 100$. (c) XT cut at $Y = 300$ (river) and at $Y = 260$ (tree).

As dynamic textures are the extension of spatial textures to the temporal do-

[1] Workshop on "An interdisciplinary approach to Textures and Natural Images Processing", Institut Henri Poincaré, January 2007, Paris, France.

main, the same difficulty for giving a proper definition arises. The description of Yves Meyer still holds, as dynamic textures are motion patterns that can be characterized by a certain repetitiveness in both space and time. Figure 1 represents two dynamic textures (downloadable at the DynTex website) where different kinds repetitiveness are observable on time and space cuts: on XT and YT cuts of the tree sequence, sine functions with or without a drift in the trend represent branches with oscillating motions due to a turbulent wind.

Given a static spatial texture, we discern three principal means whereby a static texture may become dynamic:

(1) Motion of the pattern as a whole relative to the camera (e.g. a rotating wheel, or the view of a forest from the window of a train)
(2) Change in illumination of the pattern as a whole
(3) Change intrinsic to the pattern

The first two lead the texture to become dynamic by externally imposed change to the pattern as a whole, i.e. without change in the static texture itself. The second type is rare in practice; an example would be the texture of a carpet illuminated by a flickering light bulb. In the following we refer to these first two types as weak dynamic textures.

In DynTex we are mainly concerned with the third type, and consequently refer to it simply as dynamic texture [2] . We thus understand dynamic textures as spatio-temporal patterns for which the spatial patterns are static textures that undergo pattern-intrinsic temporal change. In Section 4.3 we will further analyze the types of intrinsic changes that can occur in a dynamic texture.

3 Acquisition and Formats

3.1 Video Acquisition Protocol

As much as practically feasible, shots of dynamic textures were taken as they occur in daily life. Generally we have aimed to supply both a close-up shot of the texture, and a shot of the texture in its natural context.

The dynamic texture sequences have been acquired using a SONY 3 CCD Camera mounted on a tripod. All sequences are recorded in PAL format (720 x 576), 25 fps, interlaced. Before each shot the white balance was calibrated

[2] we use *strong*, or *strict*, dynamic texture only if we need to distinguish it from weak dynamic texture

by means of a white piece of cardboard in the ambient light conditions. The technical details are summarized in Table 1.

Those settings were determined automatically before a shot and were kept fixed during the shooting. All sequences consist of at least 250 frames, i.e 10 seconds. If sequences of standard length are required, the first 250 frames of each sequence can be used.

Acquisition equipment
Digital video camera recorder Sony DCR-TRV890E/TRV900E (3 CCD)
Tripod for stabilization
Camera Settings
Output is a PAL, IRCC norms color signal, 720 x 576, 25 fps, interlaced
Steady shot off
White balance calibrated with a white piece of cardboard
Auto but fixed focus, exposure and gain

Table 1
Acquisition equipment and camera settings

Central to the database is the so-called golden set. This set consists of high-quality dynamic texture sequences that *satisfy???* all criteria of the acquisition protocol. This dataset will stay fixed to serve as a ground truth for scientific researches. It will be possible in the future to add other dynamic texture sequences that will not be part of this golden set. See Section 4 for further details.

3.2 Video Formats

The video sequences are supplied in three formats, referred to as **raw**, **test** and **show**, respectively:

(1) **raw**: the original, DV compressed, sequences in PAL resolution (720 x 576, 25 fps). These sequences are not de-interlaced. When relevant to the texture, the avi files also contain the recorded sound track.
(2) **test**: processed sequences recommended for research testing purposes. The sequences are still DV compressed (PAL resolution), but have been de-interlaced using a spatio-temporal median filter [3]. The spatio-temporal

[3] MPlayer: http://www.mplayerhq.hu

median filter is a non-linear filter, and is simply the extension of the spatial median filter to spatio-temporal neighbourhoods. The spatio-temporal filter offers a good compromise between de-interlacing the sequence and keeping the textural information.

(3) **show**: an alternative processed format suitable for presentation purposes. The sequences are downsampled to 352x288 and encoded to DivX MPEG-4. They are of substantially lower quality than the sequences of format **test**.

4 Dynamic Texture Annotation

We have annotated the DynTex database by means of a description scheme based on the physical texture processes occurring in the sequences. The descriptors are divided over three categories [4]: *content management* descriptors, *structural* descriptors and *semantic* descriptors. The annotations serve at least three purposes: (i) they can assist users in retrieving particular dynamic textures; for example, when looking for trees in heavy wind, we may filter by selecting oscillating motions with large amplitude; or when looking for turbulent water, we select continuous texture with an irregular trajectory type and set the medium to `water`; (ii) similarly, it allows the user to quickly taylor test sets for particular research purposes. A number of tools to assist in this process are described in Section 5; and, (iii) the annotations themselves can serve as a ground truth for various problems, e.g. automatic texture characterization and texture recognition.

4.1 Dynamic Texture Modes

We found that to be able to annotate dynamic texture sequences unambiguously we need some additional terminology. First, a sequence may contain more than one dynamical texture. We refer to each of these as a dynamic texture *process*, e.g. we may have waving grass in front of waving water, giving two processes. When processes visually overlap we say that the processes *interfere*.

Texture processes are often composed of a mix of behaviors themselves. For instance, a wave may show an oscillating motion and at the same also have a turbulent foam part. To be able to annotate without having to make arbitrary choices about what is the more important or dominant behavior, we introduce the notion of a *dynamic texture mode*. [Give more and clear examples!!!] We

[4] A similar division of descriptors is used in the MPEG-7 standard

describe texture processes as a collection modes that are visually relevant in the video.

A texture mode may be either *continuous* or *discrete*. Discrete textures have directly discernable parts, e.g. a group of ants crawling around, or leaves fluttering in the wind. Continuous texture modes have continuous media, e.g. a waving flag, or are practically indiscernable from continuous media, e.g. a waving field of grass.

We will further discuss the annotation in texture modes in Section 4.3.

4.2 Content Management Descriptors

The first group of descriptors apply to the video sequence as a whole. They include administrative descriptors such as a unique identifier, location and date, and whether the protocol was followed so the sequence can be included into the DynTex golden set. If not, the deviation from the protocol is described. Next there are descriptors describing the shooting conditions, e.g. indicating if the shot was outdoor, if camera motion or a disturbance (e.g. a duck swimming into the shot) was present. Finally a number of important global properties are described: the shot type (close-up or in context), the number of dynamic texture processes visible in the sequence, and the total number of dynamic texture modes (see Section 4.3).

The complete list of management descriptors is shown in Table 2.

4.3 Structural Descriptors

Every dynamic texture mode has a unique `<DynTexId>`. If a mode is part of texture process where it is superimposed with another mode this is indicated by the `<Superposition>` descriptor. If a mode visually interferes a mode or modes from another process this is indicated by the `<Interference>` descriptor. The global texture mode descriptors are listed in Table 3. Next we turn to the visual description of the modes.

In Section 2 we saw that dynamic textures become dynamic due to intrinsic changes in the (spatial) pattern. The intrinsic changes are of two main types. By far the most common cause of change is *motion* of the constituent objects or of the medium. Motion can lead to direct change, but it can also indirectly lead to a changed appearance by the change in orientation towards light sources and camera (for instance the light reflection on water surfaces). The second, much less common cause, is by direct change of *appearance* without a causing

	General	
`<SequenceId>`	number	unique sequence identifier
`<Name>`	text	sequence name, corresponds to filename
`<Location>`	text	location (city, country) of shot
`<Date>`	text	date shot was taken
`<Protocol>`	`Golden/Extra`	flags membership golden set
`<Deviation>`	text	deviation if not member of golden set
	Acquisition conditions	
`<Outdoor>`	`true/false`	flags if shot was outdoor
`<ArtificialLight>`	`false/true`	flags if main light source artificial
`<Camera Motion>`	`false/true`	flags if camera motion present
`<Disturbance>`	`false/true`	flags if disturbance present
`<Sound>`	`false/true`	only if relevant to texture
	Global properties	
`<ShotType>`	`Closeup/Context`	type of shot
`<nProcesses>`	number	number of dynamic texture processes
`<nModes>`	number	total number of dynamic texture modes

Table 2
Content Management descriptors

motion, either by the material changing its luminant properties (e.g. flickering LEDs in a server room), or by constituent element changing their shape (e.g. a flock of birds flapping their wings).

For the annotation of the temporal dynamics, we consider we observe that whichever type of change we consider it can have the following main types: no change/still, oscillation, directed, and irregular. For motion we can work this out more precisely by means of the `<TrajectoryType>` variable with possible values: `Still`, `Oscillation`, `Directed Straight`, `Directed Curved` and `Irregular`. Similarly the `AppearanceChange` can have values `None`, `Oscillation`, `Directed` (e.g. going from one color to another, without returning to the original color), and `Irregular`.

Based on practical experience and this analysis we also found it convenient to

	General	
`<DynTexId>`	number	unique mode id
`<SequenceId>`	number	id of embedding sequence
`<Weak>`	`false/true`	flags weak dynamic texture DT
`<Discrete>`	`false/true`	flags discrete process
	Interaction Descriptors	
`<Superposition>`	`false/true`	in superposition with other mode(s)
`<Interference>`	`false/true`	interference with other process
	Semantic Descriptors	
`<MediumObject>`	text	see Section 4.4
`<Process>`	text	see Section 4.4

Table 3
Global texture mode descriptors

introduce a `<MainClass>` variable, which gives an overall characterization of the dynamic texture mode. We chose the following classes:

(1) Waving/Oscillating Motion (continuous)
(2) Directed Motion (continuous)
(3) Turbulent/irregular Motion (continuous)
(4) Oscillating Motions (discrete)
(5) Directed Motions (discrete)
(6) Irregular Motions (discrete)
(7) Direct Appearance Change

Note that the (usually discrete) "Direct Appearance Change" could be further divided by the `<AppearanceChange>` variable, but since these modes are so rare, we have collected them together in a single class.

For each dynamic texture mode the visual structure is annotated by its main type and a number of additional detailed descriptors. See Table 4.

Temporal Dynamics	
`<MainClass>`	see text
`<TrajectoryType>`	`Still/DirectedStraight/DirectedCurved /Oscillation/Irregular`
`<SpeedFrequency>`	`Low/Medium/High`
`<Amplitude>`	`Small/Medium/Large`
`<AppearanceChange>`	`None/Directed/Oscillation/Irregular`
`<TemporalRegularity>`	`Low/Medium/High`
Spatial Variation	
`<SpatialRegularity>`	`Low/Medium/High`
`<SpatialScale>`	`Fine/Medium/Coarse`
`<SpatialContrast>`	`Low/Medium/High`
`<Density>`	`Sparse/Medium/Dense`
`<DynamicsVariation>`	`Low/Medium/High`

Table 4
Texture mode structural descriptors

4.4 Semantic Descriptors

Each mode is also described with two semantic descriptors. The `<ObjectMedium>` variable is used foremost to identify the main constituent of the physical dynamic texture. To this end we must make a distinction between continuous and discrete modes. For continuous textures we provide a semantic description of the continuous *medium*. For example, for sea waves the semantic category is `water`. For discrete textures we provide category labels for the main texture *objects*, e.g. for cars on a highway the label is `car`.

For both types of textures we provide an additional semantic specifier which allows us to identify the embedding process more specifically. In the examples above, the associated specifiers are `sea` and `traffic`, respectively. Note that modes from the same process will share the same value for this variable.

The values currently used to annotate the Dyntex textures for the `<ObjectMedium>` and `<Process>` descriptors are listed below in Table 5. Note that in some cases more than one value may be applicable.

Medium Categories

textile, water, vegetation, smoke, steam, fire, cloud, foam, spray

Object Categories

droplet, branch/stem, leaf, needle, flower, car, bird, fish, flame,cloud, person

Process Categories

sea, grain, river, shower, flag, tree, shrub/plant, road, stream, waterfall, fountain, boiling, shadow, boat, aquarium, curtain, carpet, cloth, candle, sunblind, toilet, pond, source, mist, rain, escalator

Table 5
Semantic Categories

5 Querying and Browsing DynTex

The DynTex database is located at the following URL: http://www.cwi.nl/projects/dyntex/. It is freely for reserach purposes, only a registration form is required from the potential user. The sequences and their different format Golden set is the main set. An access tool is available for quickly browsing and retrieving dynamic textures of interest. Figure 2 shows this interface, where as well as a video

Figure shows different dynamic textures of the DynTex database with different numbers of processes and modes.

Here is the XML file describing the sequence 'wave' and that is available for download:


```
<Dyntex_sequence>
<SequenceId>684</SequenceId>
<Name>54pa110</Name>
<CloneId>0</CloneId>
<Date>2005-04-25T00:00:00</Date>
<Location>Amsterdam, The Netherlands</Location>
<Outdoor>1</Outdoor>
```

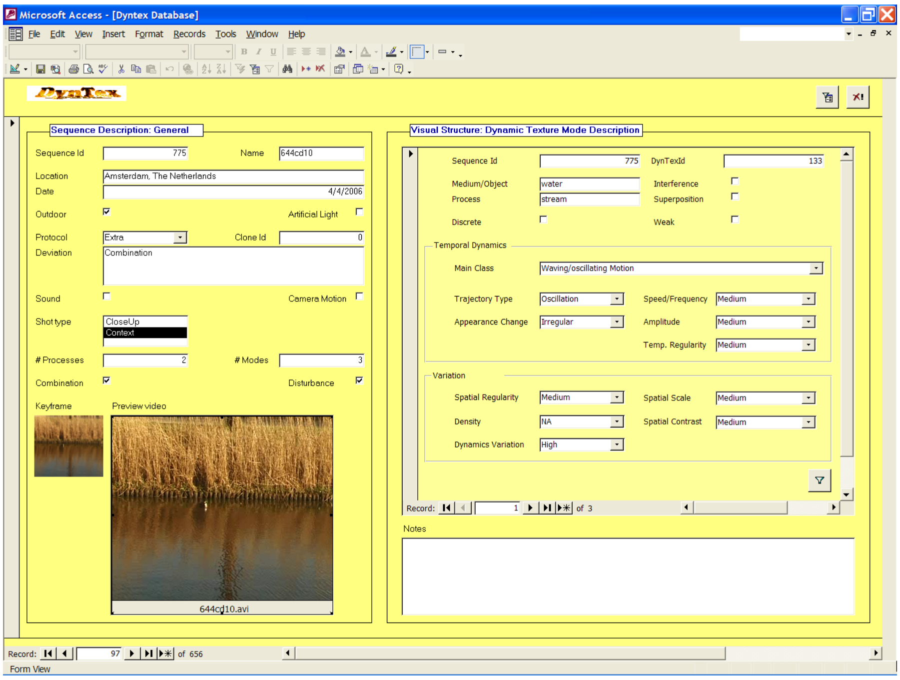


Fig. 2. Screenshot of the DynTex browser.

(a) (b) (c)

Fig. 3. (a) The *straw* sequence (1 process, 1 mode). (b) The *wave* sequence (1 process, 2 modes). (c) The *grass and river* sequence (2 processes, 1 mode each).

<ArtificialLight>0</ArtificialLight>
<Protocol>Golden</Protocol>
<Sound>0</Sound>
<CameraMotion>0</CameraMotion>
<NrOfFrames>250</NrOfFrames>
<NrOfTextures>1</NrOfTextures>
<NumberOfModes>2</NumberOfModes>
<ShotType>Closeup</ShotType>
<Combination>0</Combination>
<Disturbance>0</Disturbance>
-
<Notes>
flowers hardly move by themselves; mode 1: entire branches with flowers;
mode 2: non-flower leaves; note double semantic object!
</Notes>
<ImagePath>C:\Databases\dyntex_web\img\54pa110.png</ImagePath>

```
&lt;VideoPath&gt;C:\Databases\dyntex_web\mpeg4\54pa110.avi&lt;/VideoPath&gt;
&lt;/Dyntex_sequence&gt;
```

6 Conclusion and Future Work

Motivated by the increasing interest of the computer vision community for the study of dynamic textures and the lack of available database, we present the DynTex dynamic texture database. With more than 650 sequences of dynamic textures shot in different conditions, DynTex targets many computer vision applications (recognition of dynamic textures, spatio-temporal segmentation, synthesis. . .). It provides a set of annotated sequences that can serve for testing and comparing methods. A tool for browsing among the database as well as XML files containing the dynamic texture annotations are also made publicly available for research purposes.
Future prospects will be to define benchmarks according to specific applications (how to access and compare dynamic texture segmentation methods for instance). Additionally we are working on an relevance feedback browsing tool ([8,9]) that can search based on both the manual annotations as well as on a number of computed numerical features.

References

[1] T. Amiaz, S. Fazekas, D. Chetverikov, and N. Kiryati. Detecting regions of dynamic texture. In Lecture Notes in Computer Science, editor, *1st Int. Conf. on Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 4485, pages 848–859, 2007.

[2] A.B. Chan and N. Vasconcelos. Mixtures of dynamic textures. In *Proc. Int. Conf. Computer Vision*, pages I: 641–647, 2005.

[3] D. Chetverikov and R. Péteri. A brief survey of dynamic texture description and recognition. In *Proceedings of 4th International Conference on Computer Recognition Systems CORES'05*, "Advances in Soft Computing", pages pp. 17–26, Rydzyna, Poland, 2005. Springer-Verlag.

[4] D.Chetverikov and A.Hanbury. Finding defects in texture using regularity and local orientation. *Pattern Recognition*, 35:203–218, 2002.

[5] Yigithan Dedeoglu, B. Ugur Toreyin, Ugur Gudukbay, and A. Enis Cetin. Real-time fire and flame detection in video. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*, volume II, pages 669–673, Philadelphia, PA, March 2005.

[6] G. Doretto, D. Cremers, P. Favaro, and S. Soatto. Dynamic texture segmentation. In *Proc. Int. Conf. Computer Vision*, pages 1236–1242, 2003.

[7] J. Filip, M. Haindl, and D. Chetverikov. Fast synthesis of dynamic colour textures. In *Proceedings of the 18th IAPR Int. Conf. on Pattern Recognition (ICPR)*, pages 25–28, Hong Kong, 2006.

[8] Mark J. Huiskes. Aspect-based relevance learning for image retrieval. In W.K. Leow, editor, *Proceedings of CIVR05, LNCS 3568*, pages 639–649. Springer, 2005.

[9] Mark J. Huiskes. Image searching and browsing by active aspect-based relevance learning. In *Proceedings of CIVR06, LNCS 4071*, pages 211–220. Springer, 2006.

[10] R. Péteri and D. Chetverikov. Dynamic texture recognition using normal flow and texture regularity. In *Proceedings of 2nd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'05)*, volume 3523 of *Lecture Notes in Computer Science*, pages pp. 223–230, Estoril, Portugal, 2005. Springer.

[11] P. Saisan, G. Doretto, Ying Nian Wu, and S. Soatto. Dynamic texture recognition. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, volume 2, pages 58–63, Kauai, Hawaii, December 2001.

[12] Martin Szummer. Temporal texture modeling. Technical Report 346, MIT Media Lab Perceptual Computing, 1995. 56 pages.

[13] Martin Szummer and Rosalind W. Picard. Temporal texture modeling. In *Proc. IEEE International Conference on Image Processing*, volume 3, pages 823–826, 1996.

[14] 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 journal (TPAMI)*, 6(29):915–928, 2007.