Pattern Recognition Letters 31 (2010) 1627–1632

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

DynTex: A comprehensive database of dynamic textures ☆

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

[a] Mathematics, Image and Applications Laboratory (MIA), University of La Rochelle, avenue Michel Crépeau, 17000 La Rochelle, France
[b] MTA SZTAKI – Hungarian Academy of Sciences, 1111 Budapest, Kende u. 13-17, Hungary
[c] Leiden Institute of Advanced Computer Science (LIACS), Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 19 December 2008
Available online 13 May 2010
Communicated by H. Wechsler

Keywords:
Dynamic texture
Scientific image collections
Video segmentation
Video processing
Pattern recognition
Annotation

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. Additionally, we propose a scheme for the manual annotation of the sequences based on a detailed analysis of the physical processes underlying the dynamic textures. Using this scheme we describe the texture sequences in terms of both visual structure and semantic content. The videos and annotations are made publicly available for scientific research.

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

Just like static textures, dynamic textures can be studied in a wide variety of ways. Research on *texture synthesis* aims to model dynamic textures for realistic rendering (e.g. Szummer and Picard, 1996, Saisan et al., 2001 or Filip et al., 2006). A related area is *texture analysis* where the problem is to characterize the visual properties of the texture, e.g. its regularity or granularity. 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 (Dedeoglu et al., 2005). An interesting problem is also to discern dynamic texture from camera motion (Amiaz et al., 2007). *Texture segmentation* is about video segmentation and the accurate localization of the textures, in space, and, possibly, in time (e.g. Doretto et al., 2003; Chan and Vasconcelos, 2005). For *texture recognition* the aim is to recognize the type of dynamic texture, possibly from among several other ones (Péteri and Chetverikov, 2005; Zhao and Pietikainen, 2007); a review on DT description and recognition is presented in (Chetverikov and Péteri, 2005). In most of the cases, works on this topic consider well segmented sequences of dynamic textures. A final topic, *irregularity detection* has, to our knowledge, not been investigated so far. It aims at detecting irregular motions in sequences consisting of pure dynamic textures (for instance the detection of a piece of wood drifting on a river surface), and can be seen as the counterpart of defect detection for static textures (see Chetverikov and Hanbury, 2002).

With regard to the task of setting up test sets for performance evaluation, each of these areas of study have their own specific requirements. For instance, testing performance of methods for texture synthesis is best done on close-up sequences of the texture, whereas texture detection requires dynamic texture that is shown in 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 of the different research areas mentioned above. So far no other databases suitable for this purpose are available. One interesting pioneer database to mention is a dataset compiled at MIT (Szummer, 1995). It is composed of around 25 black and white segmented sequences of dimension $170 \times 115 \times 120$. However, this collection has a number of drawbacks: video 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

☆ http://projects.cwi.nl/dyntex/.

* Corresponding author.
E-mail addresses: renaud.peteri@univ-lr.fr (R. Péteri), sandor.fazekas@sztaki.hu (S. Fazekas), markh@liacs.nl (M.J. Huiskes).

0167-8655/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2010.05.009

show undesirable camera motion. Other datasets of dynamic textures used in research papers have been shot by the authors themselves and are not publicly released, which prevents from using them for comparison in 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 300 registered researchers (about half of which are PhD candidates) that use the database for their studies.

The structure of the paper is as follows. In Section 2 we shortly review our understanding of dynamic textures as extensions of spatial dynamic textures to the temporal domain, and discuss some implications for the compilation of the DynTex database. In Section 3 we describe the acquisition protocol of the texture sequences and the video formats in which the sequences are made available. In Section 4 we propose a video annotation scheme based on the physical processes underlying the dynamical textures. This scheme has been used for the manual annotation of the full collection of DynTex sequences. In Section 5 we discuss the two primary means by which the annotations are made available. We also discuss how users of the database can create their own test subset selections for their specific research purposes. Finally, in Section 6 we present our conclusion and discuss a number of prospects for future extensions.

2. Dynamic textures

Giving a proper definition of texture is a notoriously difficult problem. One reason is that texture has so many aspects influencing its perception as a coherent pattern. To some extent it may possess a predictable, regular organization but it will also often display a strong stochastic component. This applies to qualities such as size, orientation (or lack thereof in the case of isotropy), shape and the layout of the constituent parts. Each of these and their variations may exhibit themselves at different scales, and sometimes be scale-invariant, as in fractal patterns.

A pertinent characterization of texture was offered by the French scientist Yves Meyer who described it as "a subtle balance between repetition and innovation".[1] As dynamic textures are the extension of spatial textures to the temporal domain, the same difficulty of giving a proper definition arises there. The characterization of Yves Meyer still holds, as dynamic textures are indeed motion patterns that can be characterized by a certain repetitiveness in both space and time. Fig. 1 represents two dynamic textures (downloadable at the DynTex website) where different kinds of 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.
(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. In the following we refer to these first two types as *weak* dynamic textures.

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

For DynTex we are mainly interested in the third type, and consequently refer to this type 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 change that can occur in a dynamic texture. In the next section, we present the acquisition protocol and video formats used for creating the DynTex database.

3. Acquisition and formats

3.1. Video acquisition protocol

As much as practically feasible, the shots of dynamic textures were taken in the surroundings and circumstances in which 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 (720×576), 25 fps, interlaced. Before each shot the white balance was calibrated by means of a white piece of cardboard in the ambient light conditions. The technical details are summarized in Table 1.

The 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 s. If sequences of standard length are required, the first 250 frames of each sequence can be used.

Central to the database is the so-called *golden* set. This set consists of the texture sequences that meet the quality standards of the acquisition protocol. Specifically, they consist of sequences containing a single dynamic texture process without camera motion. This dataset will remain fixed to serve as a standard set for scientific research. In the future we will supply additional standard benchmark sets and ground truth for specific tasks; see Section 6.

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 (720×576, 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 median filter is a nonlinear filter, and is simply an extension of the spatial median filter to spatio-temporal neighbourhoods; it offers a good compromise between de-interlacing and preserving textural information.
(3) *show*: an alternative processed format suitable for presentation purposes. The sequences are downsampled to 352×288 and encoded to DivX MPEG-4. They are of substantially lower quality than the sequences of format **test**.

The next section presents how the acquired sequences have been manually annotated, based on their structural and semantic content.

[2] We use *strong* dynamic texture only if we need to distinguish it from weak dynamic texture.
[3] MPlayer: http://www.mplayerhq.hu.