



Singular Patterns in Optical Flows as Dynamic Texture Descriptors

Leandro N. Couto¹ and Celia A. Z. Barcelos²

¹ Faculdade de Computação, Universidade Federal de Uberlândia, Uberlândia, Brazil
leandronc@ufu.br

² Faculdade de Matemática, Universidade Federal de Uberlândia, Uberlândia, Brazil
celiazb@ufu.br

Abstract. This work introduces a novel approach to dynamic texture description. The proposed method is based on statistics from a vector field feature extractor that decomposes and describes features of distinctive local vector patterns as composites of singular patterns from a dictionary. The extractor is applied to a time-varying vector field, namely a dynamic texture's optical flow frames. An interest point pooling method statistically highlights the recurring texture patterns, generating a histogram signature that is descriptive of the temporal changes in the texture. The proposed descriptor is used as feature vector on classification experiments in a widespread dataset. The classification results demonstrate our method improves on the state of the art for dynamic textures with non-trivial motion, while employing a smaller feature vector.

Keywords: Dynamic textures · Optical flow · Singular patterns

1 Introduction

It is difficult to provide a formal definition of texture, but it can be described as a complex spatial arrangement of visual patterns that have particular properties and recurring characteristics [7]. Dynamic textures extend the concept of self-similarity and periodicity of textures to include repeating patterns in the temporal dimension, such as in videos that present periodic motion [2]. Dynamic texture recognition has been an area of recent interest in computer vision and pattern recognition. The most important step towards dynamic texture recognition is to determine a meaningful and discriminating computational description of the texture that is capable of capturing similarities between alike textures while distinguishing different textures. Many methods have been proposed with this goal in mind, which can be classified into four main categories: (i) motion based methods (often involving optical flow), (ii) filters and transform based methods, (iii) model based methods, and (iv) statistical methods [6]. Statistical methods are particularly effective because statistical analysis highlights the

The authors gratefully acknowledge the financial support of CNPq - National Council for Scientific and Technological Development, Brazil, Grant #309186/2017-0.

repeating local patterns that are distinctive of textures. They are among some of the best performing texture descriptor methods today, such as LBP [18] and deterministic walks [3, 6], which offer static and dynamic texture implementations. These methods describe pixel neighborhoods with efficacy and robustness to noise and changes.

Taking into account that the optical flow [14] is a suitable descriptor of motion between video frames, it has also seen extensive use as a dynamic texture descriptor [1, 5]. Statistical approaches to optical flow description have achieved good results in diverse pattern recognition tasks [12].

The work of Liu and Ribeiro [11] proposed a scale and rotation invariant feature extractor for vector fields, similar to the popular SIFT feature detector [13]. The feature extractor is based on identifying points of interest in a vector field and decomposing their region into constituent elements that are matched to a dictionary of singular vector field patterns. The method has been shown to be effective in decomposing and reconstructing flows and matching points of interest in fluid motion vector fields. Optical flows are vector fields, and as such, local regions of an optical flow can be described by their component singular patterns. Describing an entire optical flow via local patterns requires pooling the information of the local patterns into a global signature. Methods for describing a whole object by its local features are not new in computer vision; the statistical grouping of visual features can be approached in a number of ways. Many of these approaches employ histograms to compile image features, and many histogram based methods fall into the bag-of-visual-words category (also called bag-of-keypoints or bag-of-features), inspired by an analogous text analysis technique [4].

This work proposes and validates a dynamic texture descriptor based on describing the texture’s temporal aspect by employing as the feature vector a histogram of interest points acquired using the singular vector field patterns feature detector. Sections 2 and 3 describe known literature approaches to singular patterns detection and bag-of-features, respectively. Section 4 presents our novel approach to using singular patterns occurrence statistics as a global dynamic texture descriptor. The method has been validated on Dyntex, a widely used literature dynamic texture dataset. In Sect. 5, results are shown for the proposed method being applied to perform a classification task. In Sect. 6, we evaluate our proposed methods’ viability and effectiveness in comparison to other state-of-the-art techniques.

2 Singular Vector Field Patterns Detector

Most natural vector fields present particular local characteristics and points of interest. Vector field descriptors that can recognize and describe such regions have been extensively employed in several applications [10, 15]. An effective way of modeling local features of a field vector is given by Liu and Ribeiro [11]. The method is based on decomposing the vector field into singular patterns, which are components from a set of symbols, a dictionary of patterns whose

linear weighted combination is an approximation of that region of the field. An important choice is which patterns compose the dictionary, since there are no clear definitions for most visible vector field patterns of interest, such as sinks and sources [9]. Rao and Jain [16], in their seminal vector field analysis work, proposed as a dictionary six distinct patterns in which the field nullifies itself (which means the resulting vector is zero). The singular patterns detector uses a wider set of patterns as dictionary. It includes the patterns in which the flow field is nullified, but also defines a complex-valued function $F(z)$. If we choose a given point z_0 as the origin, $F(z)$ can be approximated by $f(z)$, whose Taylor expansion is a linear combination of complex basis functions, as shown in Eq. 1.

$$F(z) \approx f(z) = \sum_k a_k \phi_k(z) \quad (1)$$

In this equation, $\phi_k(z)$ are the basis flows and a_k are the coefficients that weight their effect on the local pattern around the origin, and complex basis flows are chosen instead of real ones because the former model more smooth and natural flows. The a_k coefficients are computed by cross-correlation, by projecting the vector field $F(z)$ over the basis flows. Therefore, local maxima in the sum of a_k coefficients are chosen as singular pattern points. Each pattern has an energy value which denotes its relevance. The method also builds into each chosen interest point rotation and scale invariance.

3 Bag-of-Features Based Descriptors

There are several strategies for pooling sparse local features into a signature for the whole [17]. The standard bag-of-features approach begins with the extraction of a vocabulary from a training dataset. This consists in acquiring features from diverse samples of data, clustering those features using a quantization algorithm such as k-means, and defining a vocabulary of features around the centroids of those clusters. Once that is done, new data can be described by extracting its features and clustering these new points onto the previously obtained vocabulary clusters. The number of occurrences of local features belonging to each of these clusters is a discriminating statistical aspect of the whole object; these occurrences can be captured in a histogram, which can be used as a feature vector for the complete entity. The bag-of-features histogram has great descriptive potential, given its potentially small size relative to the original data [8].

4 Singular Patterns as a Global Optical Flow Descriptor

We propose to use a set of singular patterns into a descriptor for the temporal changes in a dynamic texture, with two data pooling strategies that are customized to the specific characteristics of the singular patterns feature detector.

4.1 Best Coefficient Bag-of-features

We would usually employ a training sample to create a feature vocabulary, but, conveniently, the singular patterns detector already suggests k distinctive ϕ_k basis flows. In this case, we consider the centroids C_1, C_2, \dots, C_k in the clustering process to be given by the basis flows, such that $C_1 = (1, 0, 0, \dots, 0)$, $C_2 = (0, 1, 0, \dots, 0)$ and so forth.

Consider a new, unknown optical flow F_{new} , and its detected set of singular patterns $X_{F_{new}}$. Also consider each feature $x \in X_{F_{new}}$ is described by its corresponding a_k coefficient vector $\mathbf{a}_x = (a_1, a_2, \dots, a_k)$. For each x , its values are grouped into one of the M clusters, by the criterion of the centroid C_x closest to x . Equation 2 shows the choice of cluster for a feature x .

$$C_x = arg \min_{m|1 \leq m \leq M} |a_x - C_m|, x \in X_{F_{new}} \tag{2}$$

Once the features are assigned to clusters, a histogram $H_{F_{new}}$ of the occurrences of features in each cluster can be built to describe the set of patterns $X_{F_{new}}$. In the histogram, each bin corresponds to a C_x , and each feature $x \in X_{F_{new}}$ increments one bin, as shown in Eq. 3.

$$H_{F_{new}}(m) = \sum_{x \in X_{F_{new}}} \delta(C_x, m), \forall x \in X_{F_{new}} \tag{3}$$

Where $\delta(j, i)$ is Kronecker's delta:

$$\delta(j, i) = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases} \tag{4}$$

Notice that due to the temporal periodicity of the dynamic texture, the method is expected to detect recurring patterns, which will be highlighted by the histogram. This pooling method is shown in Fig. 1.

4.2 Coefficient Values Histogram

Another possible approach to pooling the acquired features is to consider each coefficient in the feature's vector individually. This way it is possible to measure the presence of each singular pattern component in the optical flow. To represent the occurrence of values a_k , k histograms $H_{F,k}$ are constructed, where, for each value k , the a_k coefficients for all features $x \in X_{F_{new}}$ are evaluated and organized into bins, each representing a range of values for a_k , according to Eq. 5.

$$H_{F,k}(i) = \sum_{x \in X_{F_{new}}} \mathbf{1}_A(a_x(k)), \forall x \in F, \tag{5}$$

where $\mathbf{1}_A(x)$ is the indicator function:

$$\mathbf{1}_A(x) = \begin{cases} 0, & x \notin A \\ 1, & x \in A \end{cases} \tag{6}$$

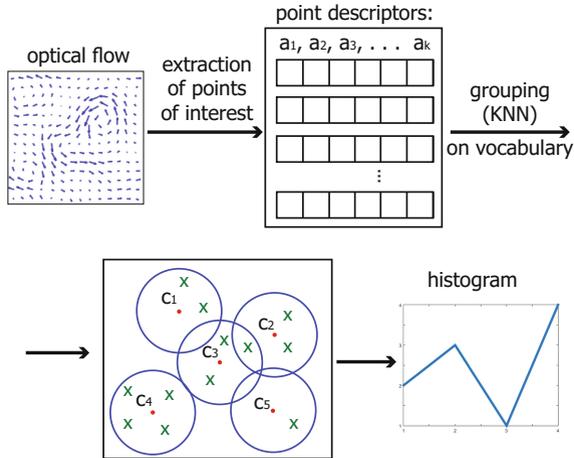


Fig. 1. Summary of the bag-of-features pooling method for vector field features.

The proposed method of pooling features components into coefficient value histograms is summarized in Fig. 2. Notice how a histogram is generated for every a_k coefficient, yielding k histograms which each describe the presence of a dictionary pattern on the composition of the optical flow.

In both methods, we append to each computed histogram its measures of energy, entropy, skewness, contrast, average, variance and kurtosis, which have been shown to significantly enhance the discriminating power of the histogram [3].

5 Experiments and Results

5.1 Experimental Parameters

The singular patterns detector employed 18 ϕ_k basis flows and pattern energy threshold in accordance with the chosen published parameters in the original work [11]. In the coefficient values histogram method, we found that 12 bins in each histogram yields good results with a reasonably sized final feature vector. One important detail is that the bins are not of equal size. We observed the coefficient value distribution resembles a Gaussian with expected value zero (coefficients can be negative) and low variance, therefore logarithmic bin sizes, with the smallest bins closest to zero, provided a much more balanced distribution of the data. Considering 18 histograms of size 12, and 7 statistics per histogram, the feature vector has 342 values.

Experiments were carried out on a challenging subset of the Dyntex dataset, with 10 different 25 frame samples (288×352 pixels) from 2 different videos for each class. Our Dyntex subset involved 79 classes, for a total 790 samples, and all classes in the dataset with non-trivial temporal variation were included.

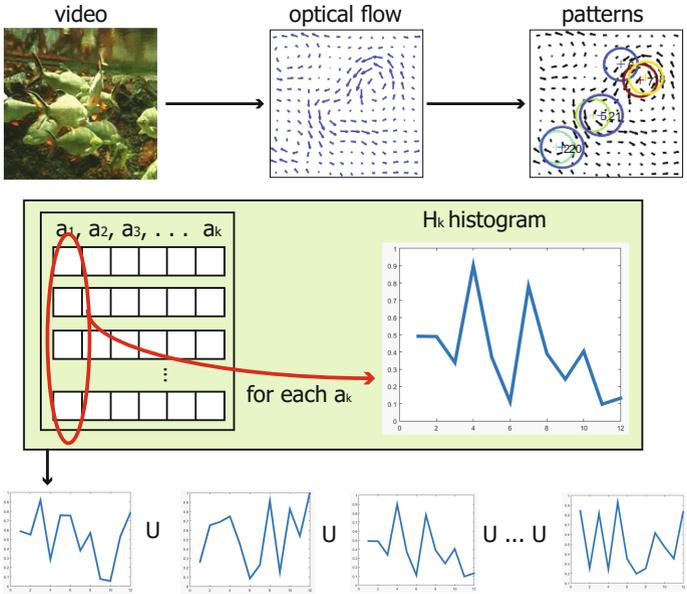


Fig. 2. Summary of the a_k coefficient values histogram generation. Here, k histograms are generated, one for each coefficient, and the feature vector is their concatenation.

Because the method has a relatively high computational cost, we also evaluated classification results after down-sampling the vector fields in half. Classification for all experiments were performed with the same parameters, using Linear Discriminant Analysis (LDA) with leave-one-out cross-validation.

5.2 Results

Table 1 presents classification results for the experiments. Results are provided with and without the inclusion of the histogram statistics metadata (referred to as *Stats* on the table), and for the statistics by themselves.

We compared our best results to the performance of LBP-TOP_[8,8,8] [18], one of the state-of-the-art dynamic texture descriptors, on the same dataset, with the same parameters and classifier, and it can be noted our descriptor outperforms LBP-TOP while using a smaller feature vector. Results for the down-sampled vector fields were also good, considering the method is about 4 times faster to compute. The option of down-sampling gives the method flexibility depending on the application’s requirements.

Table 1. Correct classification rates for singular patterns optical flow on the Dyntex dataset using both statistical pooling strategies. Spatial LBP and LBP-TOP results are also presented for comparison. Best result in bold.

Method	Dimension	Correct classifications (%)
Coefficient histogram (down-sampled)		
Stats	126	89.88
Histogram	216	84.63
Histogram + Stats	342	95.66
Coefficient histogram		
Stats	126	91.83
Histogram	216	87.80
Histogram + Stats	342	97.44
Best coefficient bag-of-features		
Stats	7	24.76
Histogram	18	44.27
Histogram + Stats	25	48.90
Coefficient hist. + bag-of-features		
Stats	133	93.17
Histogram	234	86.83
Histogram + Stats	367	97.69
LBP results for comparison		
Spatial LBP	256	87.32
LBP-TOP	768	95.98

6 Analysis and Conclusion

Our approach provides strong correct classification rates compared to LBP-TOP. Focusing on motion description alone it offers high correct classification rates for dynamic textures. When combined with a spatial descriptor like spatial LBP, the results are better than LBP-TOP, while maintaining a manageable descriptor size.

The best coefficient bag-of-features pooling approach did not fare well on its own. This is likely because most non-synthetic vector field patterns are a weighted combination of many basis flows; reducing each pattern to a single component risks discarding important information.

Realistically, vector fields are complex results of many forces. What separates one pattern from another are the a_k component coefficients. Therefore, the second approach based on values of histograms performed substantially better.

Optical flow is a strong descriptor for motion, but it is also high dimensionality data. Singular patterns address this problem, as powerful tools to summarize descriptive aspects of the flow in much less data. Our approach based on

extracted flow field features with appropriate statistical grouping yields a powerful dynamic texture descriptor of manageable dimension for most applications.

References

1. Chao, H., Gu, Y., Napolitano, M.: A survey of optical flow techniques for robotics navigation applications. *J. Intell. Rob. Syst.* **73**(1–4), 361–372 (2014)
2. Chetverikov, D., Péteri, R.: A brief survey of dynamic texture description and recognition. In: Kurzyński, M., Puchala, E., Żołnierek, A. (eds.) *Computer Recognition Systems*, vol. 30, pp. 17–26. Springer, Heidelberg (2005). <https://doi.org/10.1007/3-540-32390-2-2>
3. Couto, L., Backes, A., Barcelos, C.: Texture characterization via deterministic walks' direction histogram applied to a complex network-based image transformation. *Pattern Recogn. Lett.* **97**, 77–83 (2017)
4. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: *Workshop on statistical learning in computer vision ECCV*. vol. 1, pp. 1–2. Prague (2004)
5. Fazekas, S., Chetverikov, D.: Analysis and performance evaluation of optical flow features for dynamic texture recognition. *Sig. Process. Image Commun.* **22**(7–8), 680–691 (2007)
6. Gonçalves, W., Machado, B., Bruno, O.: A complex network approach for dynamic texture recognition. *Neurocomputing* **153**, 211–220 (2015)
7. Hájek, M.: Texture analysis for magnetic resonance imaging. *Texture Analysis Magn Resona* (2006)
8. Hoey, J., Little, J.: Bayesian clustering of optical flow fields, p. 1086. IEEE (2003)
9. Jiang, M., Machiraju, R., Thompson, D.: Detection and visualization of The Visualization Handbook 295 (2005)
10. Kihl, O., Tremblais, B., Augereau, B.: Multivariate orthogonal polynomials to extract singular points. In: *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*. pp. 857–860. IEEE (2008)
11. Liu, W., Ribeiro, E.: Detecting singular patterns in 2d vector fields using weighted laurent polynomial. *Pattern Recogn.* **45**, 3912–3925 (2012)
12. Liu, Y., Zhang, J., Yan, W., Zhao, G., Fu, X.: A main directional mean optical flow feature for spontaneous micro-expression recognition. *IEEE Trans. Affect. Comput.* **7**(4), 299–310 (2016)
13. Lowe, D.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
14. Lucas, B., Kanade, T., et al.: An iterative image registration technique with an application to stereo vision. *IJCAI* **81**, 674–679 (1981)
15. Mahbub, U., Imtiaz, H., Roy, T., Rahman, S., Ahad, A.: A template matching approach of one-shot-learning gesture recognition. *Pattern Recogn. Lett.* **34**(15), 1780–1788 (2013)
16. Rao, R., Jain, R.: Computerized flow field analysis: oriented texture fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(7), 693–709 (1992)
17. Zhang, J., Marszałek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: a comprehensive study. *Int. J. Comput. Vis.* **73**(2), 213–238 (2007)
18. Zhao, G., Pietikainen, M.: Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(6), 915–928 (2007)