Extended LBP Operator to Characterize Event-Address Representation Connectivity

Pablo Negri[1,2(✉)]

[1] CONICET, Dorrego 2290, Buenos Aires, Argentine
pnegri@uade.edu.ar
[2] Universidad Argentina de la Empresa (UADE), Lima 717, Buenos Aires, Argentine

Abstract. Address-Event Representation is a flowering technology that can change the visual perception of the computer vision world. This paper proposes a methodology to associate the input data from this kind of sensors. A new descriptor computed using an extended LBP operator seeks to characterize the connectivity of the asynchronous incoming events in a two dimensional space. Those features can be organized on histograms and combined with others descriptors, as histograms of oriented events. They can be the input of traditional classifiers to detect or recognize objects from the scene.

1 Introduction

A new paradigm for visual sensing was introduced in 2006 as the first Event-Driven Dynamic Vision Sensor (DVS) [7]. This sensor is inspired by the asynchronous Address Event Representation (AER), first introduced by Mahowald [9], and by the Kramer's transient detector concept [6]. It consists of a grid of pixels (also called "silicon retinae"), capturing changes of illumination at the focal plane. When a such event occurs, it is transmitted as an information tuple, indicating the pixel position on the grid, the time stamp and the polarity of the event. Thus, this sensor transmits a continuous flow of new events, instead of a 2D frame. This kind of vision sensor is then considered as "frameless".

The DVS visual data representation launch a new branch on the computer vision field. Traditional methodologies have to be modified in order to exploit the new information, and a new theoretical model must be developed to adapt this paradigm. Recent works tackle Visual Flow [1], Corner Detection [3] and Object Recognition [14,15] using the asynchronous temporal flow.

If the DVS camera is fixed, the images have the particularity that only moving objects are captured. In fact, the generated events correspond to the edges of this object. Static objects do not generate events. In some way, the results are similar to the Movement Feature Space [11,12] which constructs a dynamic background model using the boundaries of objects. If the temporal window for learning the background modes is fixed to some milliseconds: a moving object that stops, enters automatically to the background model. This paper seeks to generate a family of features using a histogram representation of the AER data to characterize objects shape.

The proposed descriptors are inspired on Histograms of Oriented Gradient (HOG) [4] and Local Binary Patterns (LBP) [13]. They are two of the most useful features employed nowadays on Computer Vision. The HOGs organize the gradient information in the image using histograms. The LBP operator performs a simple analysis (binary) about the relationship between the gray scale values of neighbor pixels. In detection problems, these features obtain good performances attributable to their tolerance to monotone illumination changes [10] and their simple computation. Also, they are considered as complementary features [2].

In this article, the HOG features would be arranged in a similar way as the Histogram of Oriented Level Lines (HO2L) proposed in [11]. The descriptors are organized as histograms using the orientation of the events. These orientations would be computed using the plane fitting methodology of [3]. The second contribution consists in an extended LBP operator. It seeks to characterize the connectivity of the AER data generated by the edges of the moving objects.

Next section details the steps to adapt the data provided by the DVS, followed by Sect. 3 where the histogram of oriented events and the extended LBP operator will be presented. Section 4 discusses the results and concludes the paper.

2   Features Generation

2.1   Dynamic Vision Sensor Data

A DVS reproduces the behavior of biological retinas by capturing asynchronous light changes on a 128×128 pixel grid [8]. Each change generates an event $\mathbf{e} = (p, t, pol)$, p being the spatial location on the grid, t the event time stamp, and pol defines the polarity. Polarity is a binary ON/OFF output. ON polarity captures an increase on the illumination, and OFF polarity is obtained when illumination decreases. Figure 1(a) shows the operational principle diagram of the DVS Address-Event Representation from [8].

Generally, descriptors transform visual information evaluating relationships between neighbors pixels. These associations are organized on mathematical representations as: filters, histograms, etc. Because the DVS datasets [5] available
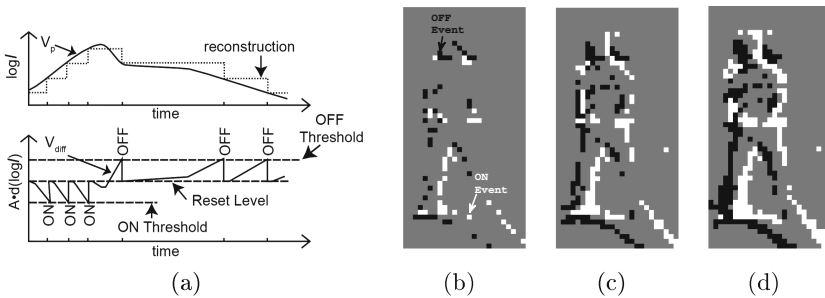

Fig. 1. (a) DVS principle of ON and OFF polarity events generation from [8], (b-d) "Street scene with cars and people walking" sample dataset from [5].

online do not provide pixels intensities or colors (new versions of DVS devices will supply this information), special methodologies should be implemented to study the pixels/event relationships.

Figures 1(b), (c) and (d), show a surveillance sample dataset downloaded from [5]. For visualization purposes the event flow composed of N events is mapped to a 2D matrix E_N which has the same size asf the retina. $E_N(p) = 1$ if the pixel p corresponds to an incoming event with ON polarity (white color). $E_N(p) = -1$ when the polarity of the event at pixel p is OFF (black color). The others pixels of E_N are set to zero (gray color). On Fig. 1(b) is represented a set (window) of $N = 100$ events. The set also corresponds to a temporal delay of 13.24 milliseconds (ms). As can be seen, the available information is not enough to recognize the person. On Fig. 1(c) and (d) $N = 300$ and $N = 500$ events, we respectively have a temporal windows of 40.7 and 67.7 ms. The number of events is now sufficient to identify a person using appropriate features.

Thus, to generate a descriptor using the DVS information, each activated pixel will evaluate its activated neighbors, within the N events on the window, and analyze their orientations and connectivity.

2.2   Events Orientation

Benosman et al. propose in [1] a regularization method to fit a plane, which considers a small spatio-temporal neighborhood $\Omega(\mathbf{e})$ around an incoming event $\mathbf{e}$. The normal component of the resulting plane is considered as a velocity vector. The direction of the normal vector defines the orientation of the event. In [3] event corners consist of those event positions where at least two valid fitting planes intersect. Clady's work uses a maximum number of events within the vicinity of the event $\mathbf{e}$. It is an interesting approach, because opening a temporal window forces to choose a fixed threshold. This threshold would not be appropriate to the dynamic of different objects in the scene. Then, events belonging to different objects/edges could be considered in the vicinity of $\mathbf{e}$ when fitting the plane to obtain the event orientation.

The orientations returned by Benosman's algorithm have values in a range of $[0, 2\pi]$. These values are converted from orientations to directions in a range between $[0, \pi]$, and discretized to integer value between $[1, V]$. This methodology maps the events from E_N to a matrix D_N.

2.3   Events Connectivity

The LBP operator was initially designed for texture recognition [13]. Let I be a gray-scale image, and the pixel $p \in I$. Thus, $LBP(p)$ assigns a label to p analyzing the gray values of their 8-connectivity neighborhood. To do this, and considering the pixels q_i a 8-connectivity neighbors of p, LBP uses an intermediate function $s(p, q_i)$ defined as:

$$s(p, q) = \begin{cases} 1, \text{ if } I(p) - I(q_i) \geq 0 \\ 0, \text{ if } I(p) - I(q_i) < 0 \end{cases} \tag{1}$$
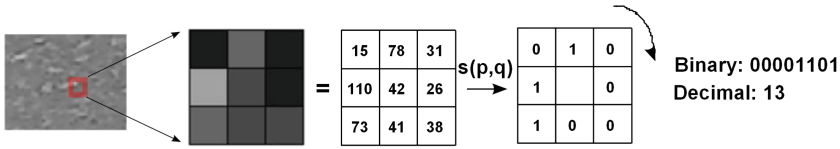
Fig. 2. *LBP* original operator computed on a *p* image pixel.

where $I(x)$ refers to the gray-scale value at position x. The label returned for the operator is obtained as: $LBP(p) = \sum_{i=0}^{8} s(p, q_i) \cdot 2^i$. Figure 2 is an example of the *LBP* computation of the $s(p, q)$ function and the output decimal label: $LBP(p) = 13$. The *LBP* operator is applied to I, giving a label to each pixel.

The patterns '0000000' (0 transition), '00110000' (2 transit.), and '11000111' (2 transit.) are considered as uniform. Others binary labels, '11011001' (4 transit.), and '01010001' (6 transit.) are not uniform. Another particularity of the *LBP* operator, is that the patterns are circular: '00110000' is the same as '11000000'. Thus, in [13] they find 9 uniform unique patterns.

To extend the LBP operator and characterize events connectivity, the Eq. 1 is modified to be adapted to the new data. The neighborhood around the central point p would be evaluated using the equality condition:

$$s(p, q) = \begin{cases} 1, \text{ if } M(p) = M(q_i) \\ 0, \text{ otherwise} \end{cases} \qquad (2)$$

The matrix $M(p)$ can be $E_N(p)$ when the polarity of the event is considered to analyze the connectivity, or $D_N(p)$ when the direction of the events is considered.

Figure 3 presents the patterns chosen to identify connectivity on DVS events and defines the operator $eLBP(p)$ around an activated pixel. Two transition patterns capture information of the extrema of a segment (patterns '18' and '21') or possible edges with more than one-pixel width ('1', '2', '3', '5', '8', '13'). Four
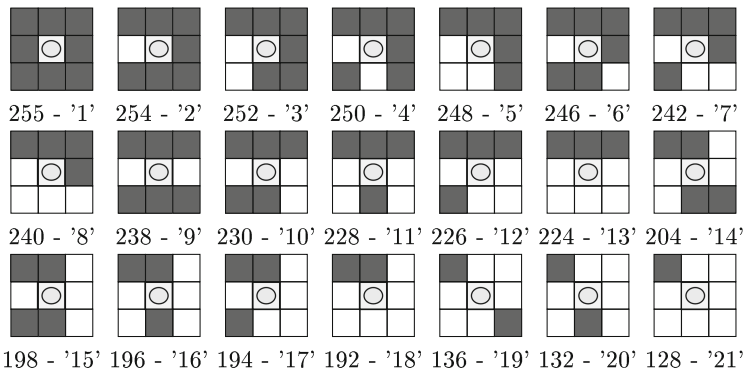

Fig. 3. Extended LBP patterns to characterize events connectivity. It shows the binary code of each pattern and the corresponding label 'x' to identify the pattern.
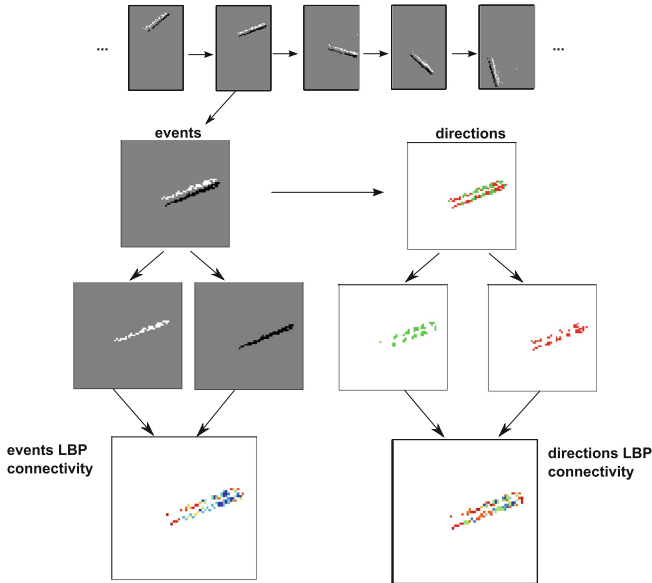
Fig. 4. Connectivity analysis using the extended LBP approach on the events generated by a rotating bar.

transition patterns characterize different configurations of possible connected edges. In total were defined 21 canonical patterns.

There are two possible analysis to study events neighborhood. Figure 4 shows an example of the AER data generated by the rotation of a bar [3]. The 2D representation was obtained histogramming events using a temporal window of $DT = 90$ ms. On the left of the figure, the extended LBP operator was applied on the events associated by their polarity. The connectivity is analyzed over the events with ON or OFF polarities. On the right side of the figure, the connectivity analysis characterizes the connectivity of the pixels with the same direction value. Each LBP label gets a specific color on the figure.

Algorithm 1 examples the computation of the extended LBP operator on the directions of the AER data. The inputs of the algorithm are the 2D matrix D

Algorithm 1. Extended LBP Computation on Events

Require: D Events Directions, C extended pattern list, V directions, $L = \oslash$
Ensure: Extended LBP matrix L.
 1: **for** $i = 1$ **to** V **do**
 2:     $\forall \, p \, / \, D(p) = i$
 3:     a) $e = eLBP(p)$
 4:     b) $c \leftarrow C[e]$
 5:     c) $L(p) = c$
 6: **return** L

---

with the direction values at each event. The directions are quantified in V values, and the list C gives the corresponding label to each LBP pattern. The output of the algorithm is the matrix L which gives to each event the eLBP label.

3   Grouping Features as Histograms

To characterize the shape of a moving object in the scene using the two dimensional features (directions and eLBP patterns), the histograms give a measure of the relationship between neighbor pixels.

Using the directions of the events, the construction of the histograms is similar to the HO2L proposed on [11]. Each bin of the histogram corresponds to one direction and accumulates the number of pixels having this value inside a region of the image. The histogram is normalized using the total number of pixels.

The histogram of the eLBP operator is computed in a similar way, and each bin corresponds to one pattern. The position of the pattern on the histogram follows the labels' numbers given on Fig. 3. The other labels which do not belong to one of the 21 canonical patterns get the label 22, and are only considered for histogram normalization by using the total number of pixels of the region.

Figure 5 shows both families of histograms computed on the Poker sequence, which was kindly provided by Dr. Bernabé Linares-Barranco [15]. The events of the sequence were collected inside a window of $N = 150$ events. Using the fitting plane algorithm, a orientation is associated to each event on E_N. Those orientations are then switched to directions $(0 - \pi)$ and quantified in 4 integer values.
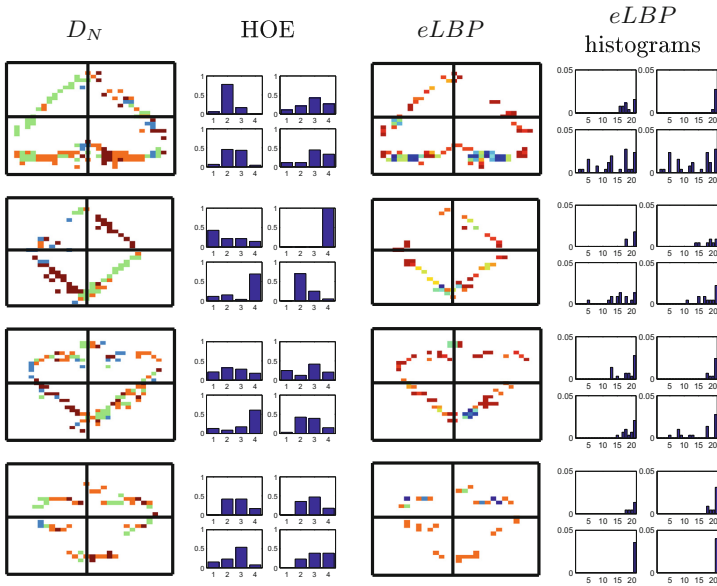

Fig. 5. HOE and histograms of extended LBP patterns obtained on the Poker Dataset.
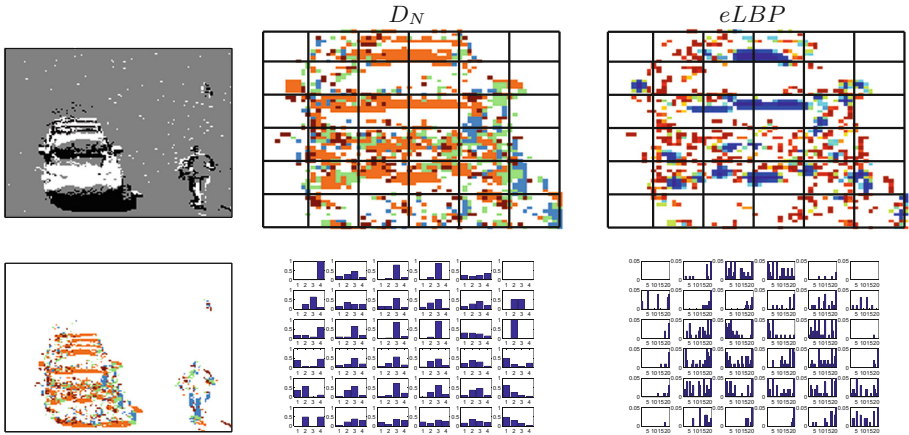
Fig. 6. Histograms analysis of the vehicle on the "Street scene with cars and people walking" dataset.

Figure 5 discriminates events directions with different colors. The image region with the card suits is divided in four patches. Inside each patch, a histogram $\mathbf{h}_i$ with V bins, where each bin counts the number of events with the associated direction. The total histogram characterizing the shape is found by concatenating them: $\mathbf{h} = \{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3, \mathbf{h}_4\}$. This representation is denominated Histograms of Oriented Events (HOE). A similar analysis is performed on the Extended LBP matrix L inside each patch obtaining the histogram $\mathbf{l} = \{\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4\}$. Figure 5 presents both features families.

It is also shown on Fig. 6 the same analysis inside an event window on the "Street scene with cars and people walking" dataset. Here, the analysis was performed on the vehicle dividing its region of interest on a grid of 6×6 patches.

4 Discussions and Conclusions

This paper proposes a histogram based feature family that shows promising discriminant properties, and can be employed later on shapes recognition.

As can be seen on Fig. 5, each HOE corresponding to a different poker sign, $\mathbf{h}^{heart}$, $\mathbf{h}^{spade}$, $\mathbf{h}^{diamant}$ and $\mathbf{h}^{clug}$, defines the shape in a discriminant way. An analysis of the eLBP histograms, will give a measure of how well the shape is defined within the events window. Uniform eLBP histograms will describe shapes with connected events. On the other hand, when the histogram has highest values at bin 21, as the *club* sign shows, the shape is not completely defined and the event window E_N has numerous isolated events.

Once the histograms $\mathbf{h}$ and $\mathbf{l}$ are obtained on a event window, they can be the input of classifiers (Boost, SVM, MLP, etc.) in order to perform the detection or recognition of a given class of object. Future research will be conducted on the implementation of shape recognition systems using both Histograms of Oriented Events (HOE) and the Extended LBP operator.

References

1. Benosman, R., Clercq, C., Lagorce, X., Ieng, S.H., Bartolozzi, C.: Event-based visual flow. IEEE Trans. Neural Netw. Learn. Syst. **25**, 407–417 (2013)
2. Cevikalp, H., Triggs, B.: Efficient object detection using cascades of nearest convex model classifiers. In: CVPR, pp. 3138–3145 (2012)
3. Clady, X., Ieng, S., Benosman, R.: Asynchronous event-based corner detection and matching. Neural Netw. **66**, 91–106 (2015)
4. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR, vol. 1, pp. 886–893 (2005)
5. JAER: Aer dataset. http://sourceforge.net/p/jaer/wiki/AERdata/. Accessed 30 May 2015
6. Kramer, J.: An integrated optical transient sensor. CS **49**(9), 612–628 (2002)
7. Lichtsteiner, P., Posch, C., Delbrück, T.: A 128 x 128 120 db 30 mW asynchronous vision sensor that responds to relative intensity change. In: ISSCC (2006)
8. Lichtsteiner, P., Posch, C., Delbruck, T.: A 128 * 128 120 db 15μs latency asynchronous temporal contrast vision sensor. JSSC **43**(2), 566–576 (2008)
9. Mahowald, M.: An Analog VLSI System for Stereoscopic Vision. Springer Science & Business Media, Heidelberg (1994)
10. Martínez-Díaz, Y., Méndez-Vázquez, H., Plasencia-Calaña, Y., García-Reyes, E.B.: Dissimilarity representations based on Multi-Block LBP for face detection. In: Alvarez, L., Mejail, M., Gomez, L., Jacobo, J. (eds.) CIARP 2012. LNCS, vol. 7441, pp. 106–113. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33275-3_13
11. Negri, P., Goussies, N., Lotito, P.: Detecting pedestrians on a movement feature space. PR **47**(1), 56–71 (2014)
12. Negri, P., Lotito, P.: Pedestrian detection using a feature space based on colored level lines. In: Alvarez, L., Mejail, M., Gomez, L., Jacobo, J. (eds.) CIARP 2012. LNCS, vol. 7441, pp. 885–892. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33275-3_109
13. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. PAMI **24**(7), 971–987 (2002)
14. Orchard, G., Meyer, C., Etienne-Cummings, R., Posch, C., Thakor, N., Benosman, R.: Hfirst: a temporal approach to object recognition. To Appear in PAMI (2015)
15. Pérez-Carrasco, J., Zhao, B., Serrano, C., Acha, B., Serrano-Gotarredona, T., Chen, S., Linares-Barranco, B.: Mapping from frame-driven to frame-free event-driven vision systems by low-rate rate coding and coincidence processing-application to feedforward convnets. PAMI **35**(11), 2706–2719 (2013)