Spike Coding for Dynamic Vision Sensors

Zhichao Bi*, Siwei Dong+,[1], Yonghong Tian+,[2], Tiejun Huang+

* *School of ECE, Shenzhen Graduate School, Peking University, Shenzhen, Guangdong, P.R. China*

+ *School of EECS, Peking University, Beijing, P.R. China*

Abstract: As an emerging kind of retinomorphic camera, the dynamic vision sensors (DVS) have shown great advantages in wide dynamic range and high temporal resolution in various applications such as autonomous driving and high-speed motion photography. However, how to compress the output spike data of DVS still remains a big challenge. To address this challenge, this paper firstly analyzes the spike firing mechanism and the redundancies of the spike data generated from DVS, and then introduces an efficient cube-based coding framework. Typically, a spike in DVS contains the location (the x-, y- addresses, the timestamp) and the polarity (On/Off). Three key strategies are designed to exploit the spatial and temporal characteristics of the spike location information for compression, including the adaptive macro-cube partitioning structure, the address-prior mode and the time-prior mode. Finally, the experimental results demonstrate that our approach achieves an impressive coding performance, with the average compression ratio of 19.519 over the original spike data, which is much higher than the results of conventional lossless coding algorithms.

1. Introduction

Conventional videos are captured and stored in the form of frames. Each frame carries the visual information from all the pixels regardless of the change of the luminance intensity, which brings significant redundancies and faces the big challenge of bandwidth limit in transmission applications. As a bio-inspired approach, the dynamic vision sensors (DVS) [1][2] have been developed, in which a three-layer retina is modelled implementing a simplified photoreceptor-bipolar-ganglion pathway [3]. Recently, the DVS devices including the dynamic and active-pixel vision sensor (DAVIS) [4], and the asynchronous time-based image sensor (ATIS) [5] are gaining more and more attentions owing to their great advantages of wide dynamic range and high temporal resolution compared to conventional image sensors in autonomous driving, high-speed motion photography, robotic automation and intelligent surveillance, etc (shown in Fig. 1).

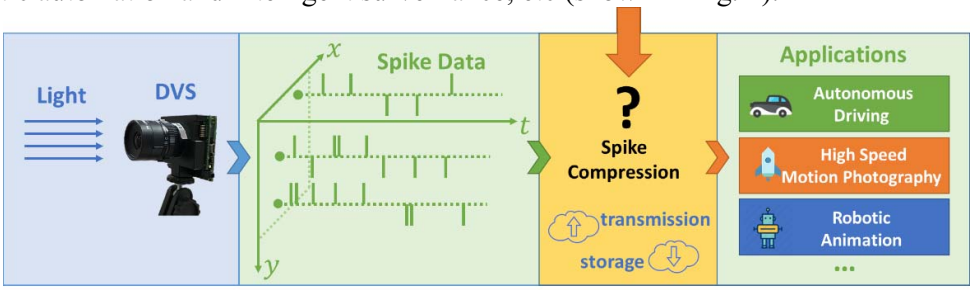

Figure 1. The perspectives and challenges of DVS.

Although these enhanced DVS devices choose different ways to represent the intensity information with extra circuits, a typical DVS sensor focuses on the luminance change and

[1] First two authors contributed equally.

[2] Corresponding author (Y. Tian, email: yhtian@pku.edu.cn).

2375-0359/18/$31.00 ©2018 IEEE
DOI 10.1109/DCC.2018.00020

does not acquire the absolute intensity for imaging in the conventional sense. The common characteristics of a typical DVS sensor can be summarized as follows. The pixels are sensitive to the scene dynamics and respond to the temporal luminance changes with spikes (or events) firing independently. Besides, the address-event representation (AER) protocol [6] is utilized to output the spikes. Moreover, the photoreceptor circuit of the pixel converts the photocurrent logarithmically to the voltage for a wide dynamic range. Generally, a spike consists of four essential elements: the x-, y-addresses, the timestamp and the polarity which can be described by a tuple $< x, y, t, p >$. The first three elements comprise the spike location in three dimensions and the last one indicates luminance change.

The brand new spike data is quite different from the conventional frame-based video sequences. A spike occurs when the luminance change reaches a predefined threshold. As a consequence, the spike data is sparse in spatial and discrete in temporal. With the AER protocol, dynamic vision sensors can be built with a rather high time resolution of 1 μs (DAVIS 240B). The timestamp of each spike consumes the most bits. For example, in DAVIS 240B, the timestamp in represented with 32 bits while a spike is measured with 64 bits. Meanwhile, the spatial resolution is limited by the output bandwidth because of the bit cost of the x- and y-addresses. Thus, the coding for the spike data of dynamic vision sensors is a challenge demanding prompt solutions.

The rest of the paper is organized as follows. In Section 2, the conventional video coding techniques and the spike coding in the field of neural computation are reviewed. In Section 3, we investigate the characteristics of the spike data. Then, a cube-based coding framework to compress the spike location and its polarity is described in greater detail in Section 4. Next, a DVS spike dataset is introduced in Section 5 along with the demonstration of experimental results. Finally, we conclude the paper in Section 6.

2. Related Works

Conventional video coding technologies have been researched and developed in about four decades, leading to several major standards including MPEG-4 AVC/H.264, MPEG-HEVC/H.265, AVS and VP9, etc [7-10]. Among them, the block-based hybrid coding framework including motion compensation, transform, quantization and entropy coding has efficiently reduced the redundancies in the video data. However, compared to the frame-based videos, the spike data generated from DVS have different properties. It is not straightforward to apply conventional video coding techniques to the spike coding.

The spike firing mechanism in dynamic vision sensors is inspired from the discharge activity of neurons. There are a number of researches trying to estimate the entropy or the coding capacity of the spike train in the field of neural computation. Mackay *et al.* [11] initially applied entropies into neural coding regarding the spikes as binary strings. Strong *et al.* [12] proposed the "direct method" in which the spike train is discretized into binary words with a particular letter size. Furthermore, the work proposed in [13][14] indicates that the inter-spike interval can be modelled by a suitable probability density function. The entropy can be estimated via numeric methods at any desired temporal resolution [14].

These explorations are from the point of view of a neuron based on biological models. Even though the rapid developing semiconductor technology enables the implementation of intelligent sensors to be so close to the biological neural systems, some approaches and

structures taken by nature are not able to be applied in a feasible way, such as the parallels in neural signaling and the complicated connections between neurons. Hence, we mainly focus on the spike coding for practical vision sensors and absorb the edifying experiences from the predecessors in neural computation as well as in conventional video coding.

3. Spike Firing Mechanism and Analysis

In this part, we will look into the spike firing mechanism of DVS from the microscopic analysis. The photoreceptor in the pixel captures the light and logarithmically converts the photocurrent into a voltage output [2]. A spike is fired when the difference of the luminance intensities (measured by voltage) reaches a predefined threshold ϕ:

$$\left|logI_{ph} - logI'_{ph}\right| > \phi \tag{1}$$

where I_{ph} and I'_{ph}denote the photocurrent at current moment t and at the moment $t'$when the last spike was elicited in the same pixel, respectively. In order to analyze the spike generation, we should look into the luminance intensity, which is often considered to be proportional to the number of the recorded photons $R(t,\delta)$ during the time interval $(\mathrm{t},\mathrm{t}+\delta]$. Also, we define $N(t,\delta)$ as the number of photons arriving during the same time interval $(t,t+\delta]$. It is generally assumed that $N(t,\delta)$ is a homogeneous Poisson process in which the photon arrival rate λ is constant in a short period of time. Thus the probability of n photon arrivals in any time interval of length δ is given by:

$$\Pr\{N(t,\delta)=n\} = e^{-\lambda\delta}\frac{(\lambda\delta)^n}{n!} \tag{2}$$

In ideal circumstances, all of the photon arrivals can be recorded in the photoreceptor, i.e. $R(t,\delta)=N(t,\delta)$. However, considering the presence of the dead time between two consecutive photon arrival events, if the former event is recorded at t, any latter event arriving during the time interval $(t,t+\tau]$ will not be recorded, where τ (usually $\tau<\delta$) is the length of the dead time. Hence, $R(t,\delta)$ is non-Poisson distributed [15].

As $N(t,\delta)$ is a Poisson process, the time difference Y of two consecutive photon arrivals, also known as inter-arrival, is exponentially distributed with mean $1/\lambda$. With the presence of the dead time τ, the time difference of two consecutive recorded photon arrivals Y_ϵ can be described as a shifted exponential distribution with mean $\tau+1/\lambda$:

$$Y_\epsilon = Y + \tau \tag{3}$$

And the corresponding probability density function starting at τ would be as follows:

$$f(y_\epsilon) = \begin{cases} \lambda e^{-\lambda(y_\epsilon-\tau)}, & y_\epsilon \geq \tau \\ 0, & y_\epsilon < \tau \end{cases} \tag{4}$$

From the above, though $R(t,\delta)$ is a non-Poisson process, it can be modeled as a renewal process, in which the inter-arrival intervals are positive, independent and identically distributed (IID) random variables. In this case, the intervals between one recorded photon and the next are independent and shifted exponential distributed. Thus, the photoreceptor returns to a state probabilistically equivalent to the starting state over at each recorded

photon arrival epoch. Because of the renewal property, each photon record can be regarded as a renewal ϵ [16].

Therefore, we can derive the probability density of $R(t,\delta)$. Since the expression of at least n photon arrivals recorded in $(t, t+\delta]$, is equivalent to that the time interval w_n of n recorded photon arrivals is no longer than t, we have

$$\begin{aligned}\Pr\{R(t,\delta)=n\} &= \Pr\{R(t,\delta)\geq n\} - \Pr\{R(t,\delta)\geq n+1\} \\ &= \Pr\left\{\sum_{i=1}^{n} y_\epsilon \leq \delta\right\} - \Pr\left\{\sum_{i=1}^{n+1} y_\epsilon \leq \delta\right\} \\ &= \Pr\{w_n \leq \delta\} - \Pr\{w_{n+1} \leq \delta\}\end{aligned} \tag{5}$$

where $w_n = \sum_{i=1}^{n} Y_\epsilon$ denotes the time interval of n recorded photon arrivals. It is easy to derive that w_n is generalized gamma distributed with parameters λ, n and τ.

$$f(w_n;\lambda,n,\tau) = \frac{(w_n-n\tau)^{n-1}}{\Gamma(n)}\lambda^n e^{-\lambda(w_n-n\tau)}; \quad w_n \geq n\tau. \tag{6}$$

Thus we have

$$\begin{aligned}\Pr\{R(t,\delta)=n\} &= F_n(\delta) - F_{n+1}(\delta) \\ &= \frac{\gamma\big(n,\lambda(\delta-n\tau)\big)}{\Gamma(n)} - \frac{\gamma\big(n+1,\lambda(\delta-n\tau-\tau)\big)}{\Gamma(n+1)}\end{aligned} \tag{7}$$

where $F_n(\cdot)$ and $\gamma(\cdot)$ denote the cumulative distribution function of w_n, and the lower incomplete gamma function, respectively.

We use a video captured by a high speed camera to validate our model. Since the pictures are shot within a short period of time, the photon arrival rate λ is assumed constant. The values of a certain pixel is able to represent the luminance intensity (proportional to recorded photon arrivals $R(t,\delta)$) during $(t, t+\delta]$ where t is the starting time of each shot, and δ here is the exposure time. We plot the histogram of the values of a certain pixel, and fit the density function $R(t,\delta)$. Fig.2 shows $R(t,\delta)$ well fits the histogram.

Due to $\tau \ll \delta$ in practical applications, $R(t,\delta)$ is approximately normal distributed with mean $E[R(t,\delta)]$ and variance $Var[R(t,\delta)]$. Fig.2 shows that the approximate normal distribution also fits the pixel value histogram well.

$$E[R(t,\delta)] \sim \frac{\lambda\delta}{\lambda\tau+1}, \quad Var[R(t,\delta)] \sim \frac{\lambda\delta}{(\lambda\tau+1)^3} \tag{8}$$

Once $R(t,\delta)$ is approximately normal distributed, Eq.1 can be rewritten as

$$\left|\log\frac{I_{ph}}{I'_{ph}}\right| = \left|\log\frac{R(t,\delta)}{R(t',\delta)}\right| = |\log Z(t,t',\delta)| > \phi. \tag{9}$$

where Z is the ratio of two independent normal random variables $R(t,\delta)$ and $R(t',\delta)$.

The ratio Z can be linearly transformed as follows [17].

$$Z \sim \frac{\sqrt{\lambda\delta(1+\lambda\tau)}+x}{\sqrt{\lambda'\delta(1+\lambda'\tau)}+y} \tag{10}$$

where x and y are independent standard normal random variables, and the photon arrival rate λ and λ' are assumed to be constant during $(t, t+\delta]$ and $(t', t'+\delta)$, respectively.

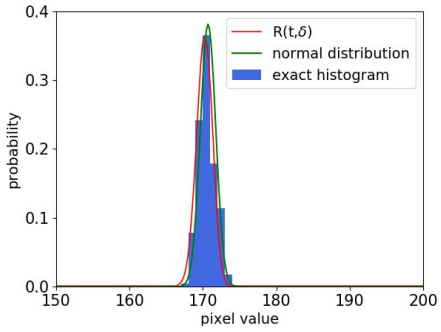

Figure 2. Histogram of the values of a certain pixel (in blue), the fitted density function $R(t,\delta)$ (red line), and the approximate normal distribution (green line).

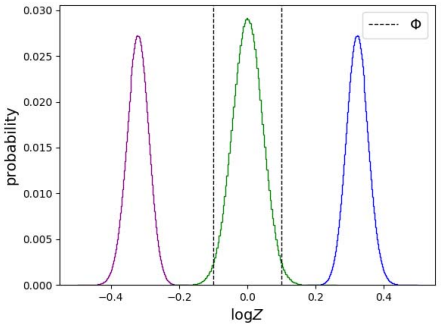

Figure 3. The distribution of $\log Z$.
Green line: when $R(t,\delta)$ and $R(t',\delta)$ are approximately identically distributed.
Purple and blue lines: when $R(t,\delta)$ and $R(t',\delta)$ have different probability densities.
ϕ: the spike firing threshold in DVS.

Since Eq.10 is quite complicated, we focus on qualitative analysis of the density distribution, such as the shape of the density of Z which, to a certain extent, determines the generation of spike in DVS. Z has either a unimodal or a bimodal density. For many practical situations, the densities which are theoretically bimodal, can be approximate to unimodal [17]. Considering the monotonically increasing of the logarithm function, the density shape of $\log Z$ is similar to that of Z.

After a spike was just elicited at the moment t', $R(t,\delta)$ is similar to $R(t',\delta)$, $\log Z$ is approximately symmetric around the ordinate origin (the green line in Fig.3), which indicates the probability of firing a new spike $\Pr\{|\log Z(t,t',\delta)| > \phi\}$ is infinitesimal. Once the luminance intensity variation causes significant differences between $R(t,\delta)$ and $R(t',\delta)$, the changed distribution of $\log Z$ (the purple or the blue line in Fig 3), is more likely to satisfy the spike firing condition. When a new spike is fired, t' is reassigned to the latest firing moment and the time distance between t and t' returns to be close, then the variation of $\log Z$ starts again from scratch.

Based on the analysis above, we conclude that,

- The process between two successive spikes at a certain pixel can be described with a specific probability model, such as a renewal process. For a certain pixel, only when the probability density is significantly changed, a spike is elicited. Thus, the spike generation mechanism of DVS has taken advantage of reducing the temporal redundancies.
- When the luminance intensity is stably changing, such as in linear increasing or decreasing, the time intervals between consecutive spikes at the same pixel appear to be nearly equal, which may induce temporal correlations.

- Besides, considering that the pixels are independently recorded and the adjacent pixels receive almost the same luminance intensities simultaneously, the spatial redundancies still exist.

In section 4, we design several strategies and concentrate on removing the redundancies.

4. The Spike Coding Framework

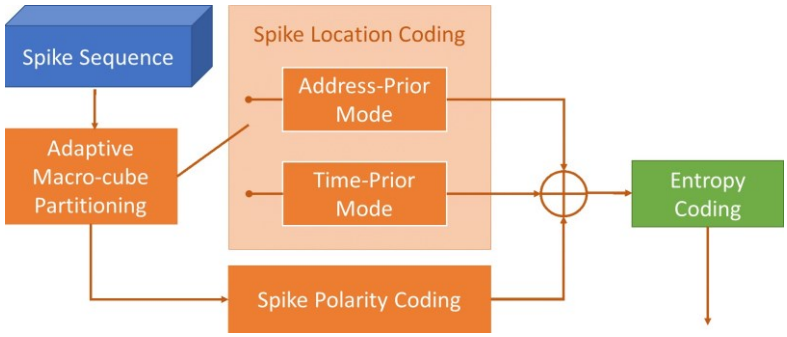

Figure 4. The spike coding framework

To exploit the temporal and spatial redundancies analyzed in Section 3, a cube-based spike coding framework (shown in Fig. 4) is proposed which proceed as follows.

A DVS spike sequence can be partitioned into multiple macro-cubes in temporal, each of which has the full spatial resolution of the pixel array. A macro-cube is then spatially split into small spike-cubes (shown in Fig. 6). The encoding process for the spike-cube consists of the spike location coding and the spike polarity coding. The former is performed with the optimization decision of the address-prior (AP) mode and the time-prior (TP) mode, while the latter utilizes the previous coding information as the reference to predict the imminent polarity. The prediction residuals in AP mode, TP mode and the polarity coding are fed into an adaptive context-based entropy coder. We describe these in more details in the following sections.

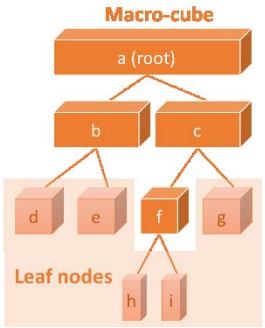

Figure 5. The adaptive macro-cube partitioning structure.

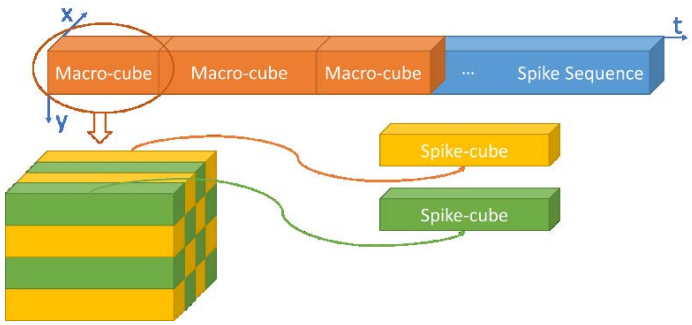

Figure 6. A macro-cube is divided into small spike-cubes.

4.1. Adaptive Macro-cube Partitioning

In practical, the spike numbers are highly volatile in the fixed time period, associating with the luminance change and objects' motion. To smooth the fluctuation and to limit the context size of the entropy coder, the macro-cubes shall be adaptively allocated with approximately equal number of spikes.

At the first stage, the macro-cube is initially set to the maximum time-length L_{max} , as the root of a binary tree. Starting from the root, if the total number of the spikes in each node is smaller than a predefined threshold, it shall be equally split into two small macro-cubes recursively. The leaf nodes of the binary tree form the final shape of the macro-cube partitioning. Each node is encoded with a split flag identifying that whether current macro-cube is about to be split or not. Fig. 5 demonstrates the partitioning process.

4.2. Spike Location Coding

The spike distribution in spatial can be classified into two types: spatial-centralized and spatial-decentralized. Objects' movements usually lead to a number of adjacent pixels to elicit spikes (spatial-centralized). Moreover, the luminance variation would cause spike firing globally (spatial-decentralized). Thus, the time-prior (TP) mode and the address-prior (AP) mode are designed for these two types, respectively. Each mode will be tried and the best one with minimum rate cost is chosen.

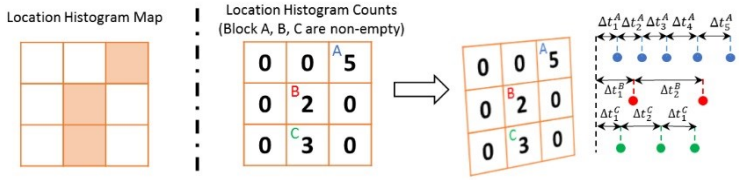

Figure 7. The location histogram map (orange blocks refer to non-empty blocks) and the location histogram counts.

A. Address-Prior Mode

The AP mode is designed for the spatial-decentralized spike-cubes. By projecting all the spikes in a spike-cube to the xy plane, a location histogram is depicted [18]. We represent the location histogram using location histogram map and location histogram counts. The former is a binary map indicating whether the spikes exist or not, while the latter records the number of spikes at each pixel.

Similar to the location histogram map coding in [18], we code the blocks utilizing a context-based arithmetic coder and use the information of coded blocks in neighborhood as references to adaptively update the context. Moreover, the corresponding blocks of the histogram map in the previous spike-cube are also used as references to help adjust the context for current block. The non-empty blocks of the location histogram counts are encoded as individual symbols.

In general, the luminance changes smoothly. The time intervals between consecutive spikes for a certain pixel show the temporal relevance which can be used to predict the occurring time of the subsequent spike. In this way, the timestamps of the spikes fired at each pixel are differential coded one after another (shown in Fig. 7).

B. Time-Prior Mode

To code the spatial-centralized spike-cubes, the time-prior mode is an appropriate solution. We first find a center point (x_c^*, y_c^*) and project all the other spikes to its time axis which appears to be a timeline. The center point is determined by:

$$x_c^*, y_c^* = arg \min_{x_c, y_c} |x_i - x_c| + |y_i - y_c| \tag{12}$$

where (x_i, y_i) denotes the spike address in spatial. The center point (x_c^*, y_c^*) is either encoded directly or differential encoded by referencing the center point $(x_{c,prev}^*, y_{c,prev}^*)$ of the previous collocated spike cube.

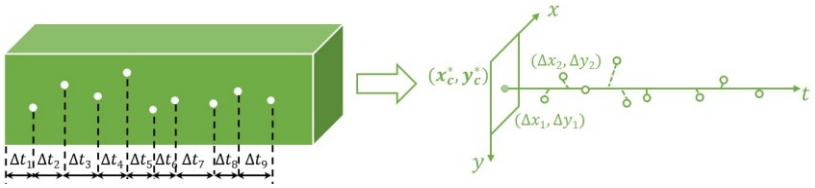

Figure 8. The motion vectors of the spikes to the center and the differences of timestamps are encoded in TP mode.

As the spikes are centralized in spatial, the offsets of the spatial location (x_i, y_i) to the center point (x_c^*, y_c^*) form a motion vector $(\Delta x_i, \Delta y_i)$ which is fed to the entropy coder. The timestamps of the spikes in the center point's timeline are also differential encoded like that in AP mode (shown in Fig. 8).

4.3. Spike Polarity Coding

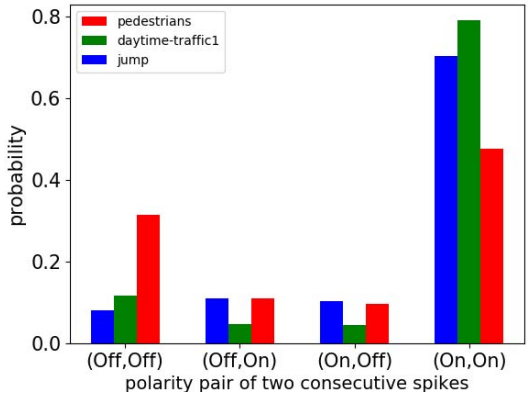

Figure 9. The temporal correlation of the polarity pairs of two consecutive spikes (statistics on three different DVS sequences).

In DVS, the spike polarity indicates the luminance change, while "On" means luminance increase and "Off" means decrease [2]. By analyzing the polarity pair of two consecutive spikes in three DVS sequences, the temporal correlation of the polarity is relatively high. In Fig. 9, if the previous spike contains an "On" (or "Off") polarity, there is a large probability that the following spike will have the same polarity of "On" (or "Off"). This is mainly because that in general, the luminance is increasing or decreasing in a steady state. Thus, the encoding of the spike polarity exploits the previous spike polarity as the context for the current one.

5. Experiments

To evaluate the performance of the proposed coding framework, we constructed the PKU-DVS dataset for spike coding [19]. The PKU-DVS dataset is classified into two categories (Class A for indoor and Class B for outdoor scenarios), which are captured under various circumstances including daytime, night, near view, distant view and high speed movements.

Table 1 demonstrates the results of the coding experiments on the PKU-DVS dataset and we can find that compared to the original spike data, the average compression ratios are

21.482 for Class A, 17.230 for Class B, and 19.519 for the whole. In contrast, the conventional lossless coding algorithms only achieve average compression ratios of 4.407 (Lempel-Ziv compression algorithm, LZ77) [20] and 12.642 (Lempel-Ziv-Markov chain algorithm, LZMA) [21] in total, respectively. Notably, for high speed motion scenarios with high spike firing rates, such as the sequence waterdrop, fluorescent and lighter, our approach manifests an impressive coding performance. For night scenarios, due to the influence of noise, the instinct temporal correlation is disturbed which may slightly lower the coding performance. In general, the proposed framework is still capable of coding the spike data efficiently.

Table 1. Performance evaluation of the proposed coding framework and comparison against conventional lossless coding algorithms (each spike is 64 bits).

Sequence		Time (s)	Total spike number	Compression ratio (compared to the original data)		
				Our approach	LZ77	LZMA
Class A: indoor	waterdrop	3.80	11563244	59.594	13.364	52.953
	fluorescent	5.44	11643175	33.677	4.025	14.669
	lighter	2.10	2792140	20.383	4.458	12.644
	football	7.21	9745102	18.816	5.935	14.485
	jump	3.28	2375023	8.967	3.383	7.351
	game	9.57	5918278	4.496	2.857	5.223
	pendulum	5.37	113683	4.438	2.497	4.078
	Average for Class A	**5.25**	**6307235**	**21.482**	**5.217**	**15.915**
Class B: outdoor	intersection	10.35	30483325	62.198	4.282	18.555
	pedestrians	355.21	25455049	14.247	4.323	10.430
	daytime-traffic1	109.84	14246080	11.132	3.765	8.630
	daytime-traffic2	301.60	5525454	6.487	2.542	4.573
	night-roadside	63.37	17018998	5.441	3.178	5.959
	night-traffic	15.47	5423636	3.874	2.687	4.798
	Average for Class B	**142.64**	**16358757**	**17.230**	**3.463**	**8.824**
Total average		**68.66**	**10946399**	**19.519**	**4.407**	**12.642**

6. Conclusion

In this paper, to compress the spike data generated from the dynamic vision sensors, we deeply investigated the spatial and temporal correlations of the spikes, and a lossless cube-based spike coding framework was proposed. The spike data including the spike location and the polarity are efficiently encoded with several strategies: the adaptive macro-cube partitioning structure, the address-prior and the time-prior modes. The experimental results over the PKU-DVS dataset show that our approach has achieved an impressive coding performance with the average compression ratio of 19.519 against the original data. In the future work, we will investigate more coding strategies to enhance current framework. Meanwhile, the lossy coding methods for the DVS spikes are also worth studying.

Acknowledgements. This work is partially supported by grants from the National Basic Research Program of China under grant 2015CB351806, the National Natural Science Foundation of China under contract No. 61390515, No. U1611461, and No. 61425025, and also supported by Beijing Municipal Commission of Science and Technology under contract Z171100000117008.

References

[1] Lichtsteiner P, Delbruck T. A 64x64 AER logarithmic temporal derivative silicon retina. Research in Microelectronics and Electronics, 2005. IEEE, 2005, 2: 202-205.
[2] Lichtsteiner P, Posch C, Delbruck T. A 128x128 120dB 15us latency asynchronous temporal contrast vision sensor[J]. IEEE Journal of Solid-State Circuits, 2008, 43(2):566-576.
[3] Posch C, Serrano-Gotarredona T, Linares-Barranco B, et al. Retinomorphic event-based vision sensors: bioinspired cameras with spiking output. Proceedings of the IEEE, 2014, 102(10): 1470-1484.
[4] Brandli C, Berner R, Yang M, et al. A 240× 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor. IEEE Journal of Solid-State Circuits, 2014, 49(10): 2333-2341.
[5] Posch C, Matolin D, Wohlgenannt R. An asynchronous time-based image sensor. Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on. IEEE, 2008: 2130-2133.
[6] Boahen K A. Point-to-point connectivity between neuromorphic chips using address events. IEEE Transactions on Circuits & Systems II Analog & Digital Signal Processing, 2000, 47(5):416-434.
[7] Wiegand T, Sullivan G J, Bjontegaard G, et al. Overview of the H. 264/AVC video coding standard. IEEE Transactions on circuits and systems for video technology, 2003, 13(7): 560-576.
[8] Sullivan G J, Ohm J, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard. IEEE Transactions on circuits and systems for video technology, 2012, 22(12): 1649-1668.
[9] Yu L, Chen S, Wang J. Overview of AVS-video coding standards. Signal processing: Image communication, 2009, 24(4): 247-262.
[10] Mukherjee D, Bankoski J, Grange A, et al. The latest open-source video codec VP9-an overview and preliminary results. Picture Coding Symposium (PCS), 2013. IEEE, 2013: 390-393.
[11] MacKay D M, McCulloch W S. The limiting information capacity of a neuronal link. Bulletin of Mathematical Biology, 1952, 14(2): 127-135.
[12] Strong S P, Koberle R, van Steveninck R R R, et al. Entropy and information in neural spike trains. Physical review letters, 1998, 80(1): 197.
[13] Victor J D. Binless strategies for estimation of information from neural data. Physical Review E, 2002, 66(5): 051903.
[14] Reeke G N, Coop A D. Estimating the temporal interval entropy of neuronal discharge. Neural computation, 2004, 16(5): 941-970.
[15] Daniel F Y, Fessler J A. Mean and variance of single photon counting with deadtime. Physics in medicine and biology, 2000, 45(7): 2043.
[16] Feller W. An introduction to probability theory and its applications. John Wiley & Sons, 2008.
[17] Marsaglia G. Ratios of normal variables. Journal of Statistical Software. 2006.
[18] Tsai S S, Chen D, Takacs G, et al. Improved coding for image feature location information. Proc. SPIE. 2012, 8499: 1-10.
[19] PKU-DVS dataset: http://pkuml.org/resources/pku-dvs.html
[20] Ziv, Jacob, and Abraham Lempel. "A universal algorithm for sequential data compression." IEEE Transactions on information theory 23.3 (1977): 337-343.
[21] Pavlov, Igor. "LZMA SDK (software development kit)." (2013).