Proceedings of the 4th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/49/20241194

Research on the analog-to-digital hybrid square wave encoder signal error compensator

Shouhuan Zhang

School of Astronautics, Harbin Institute of Technology, Harbin, 150001, People's Republic of China

1581047570@qq. com

Abstract. Photoelectric encoders are important electronic components, widely used in motion control systems, position measurement systems, and can measure angular displacement or displacement with high accuracy. Encoders widely used today output analog sine waves or digital square waves. The resolution of the encoder determines the measurement accuracy of the encoder, because the sine wave encoder can be more easily subdivided, which gradually replaces the square wave encoder in many occasions, but the square wave encoder is still widely used in the system with low demand for accuracy and low cost. The prerequisite for high subdivision of the output of the square wave encoder is good signal quality. Through the study of the signal output by the ideal encoder, this paper proposes the characteristics of distinguishing the error signal from the ideal signal, divides the error into two categories, designs a filter with two steps to compensate separately, and uses Simulink to establish a simulation model for verification.

Keywords: Encoder, Square Wave, Signal Processing, Error Compensation.

1. Introduction

The measurement of diagonal displacement is often essential for closed-loop servo motion control systems. Photoelectric encoders are the most common sensors for measuring displacement or angular displacement, which converts displacement or angular displacement into an electrical signal through an optoelectronic process[1][2]. The optical encoder is composed of a light source, an optical code disc and a photosensitive element. Taking a rotary photoelectric encoder as an example, the coding disk is a certain diameter disk with a large number of rectangular clear apertures. By connecting the coding disk with the rotation axis coaxially, the code disc is driven to rotate. Whenever the code disc rotates to the position of the photosensitive element, the clear aperture, and the light source in the same straight line, the photosensitive element will output an electrical signal, thereby determining the angluar position of the code disc.

It can be simply seen that by increasing the number of rectangular clear holes on the code disc, the number of pulses generated per rotation (Pulse Per Revolution, PPR) increases, and therefore the resolution of the encoder increases. However, due to process and cost constraints, it is impossible to increase PPR indefinitely. By utilizing the rising and falling edges of two quadrature digital signals, a four-fold subdivision of digital encoders can be achieved. To improve the resolution of an incremental sin-cos quadrature analog encoder, there are more methods available, such as amplitude division,

© 2024 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/49/20241194

frequency multiplication, or math functions [3–5]. But there are fewer studies of more than four times the subdivision methods for digital output.

Ming-Shyan Wang, etc., proposes a method that allows a 64-fold subdivision of square wave signals using FIPS (Frequency Independent Phase Shifter) [6]. However, due to poor signal quality, this approach can lead to problems such as lost pulses.

In fact, the actual validity of the subdivision signal obtained by high subdivision of poor quality output signals is questionable. The quality of high-power subdivision signals can be ensured by analyzing and compensating the sources of errors to improve signal quality. We are able to analyze and compensate for error sources to improve signal quality by means such as Lisa Ru graphics, which ensures the reliability of the high magnification sub-divsion of sine and cosine encoders.

In this paper, the error factors that may cause the quality of the square wave output to deteriorate are analyzed, and the characteristics of distinguishing the error signal from the effective signal are given, and a method is designed to improve the signal quality of the square wave output.

2. Signal characteristics

At a constant speed, an ideal digital encoder would output two quadrature square wave signals with a duty cycle of 50%. We show that although the time of the high and low levels of the encoder output continues to change during acceleration or deceleration, the times of the high and low levels can be considered approximately equal over a period of the square wave.

Suppose a high/low corresponds to an angle of θ, a high/low duration is T, and the angular velocity is ω,there is obviously:

$$\frac{2\pi}{2*\mathrm{PPR}} = \theta \tag{1}$$

$$\frac{\theta}{\omega} = \mathrm{T} \tag{2}$$

Since the time T is short, we can approximate the duration of the next low/high level as $\mathrm{T} + \Delta\mathrm{t}$.

$$\mathrm{T} + \Delta\mathrm{t} = \frac{\theta}{\omega + \Delta\omega} \tag{3}$$

$$\Delta\omega = \ddot{\theta}\mathrm{T} \tag{4}$$

$$\frac{\Delta\mathrm{t}}{\mathrm{T}} = -\frac{\theta}{\omega^2}\ddot{\theta} \tag{5}$$

For the 4096-line encoder, it has been calculated that even at lower speeds (100r/min) with large acceleration ($1000\mathrm{rad}/\mathrm{s}^2$), the difference in adjacent high and low level duration is only 0.75%. Therefore, it is reasonable to assume that the signal with a large occasional duration difference between high and low levels is caused by error. It is worth noting that when the encoder is running at very low speeds, the assumption that the output square wave duty cycle should be 50% no longer holds, so this processing method will fail.

Examples include accidental errors in the position and width of the aperture on the code disk, unexpected pulses caused by noise, and time imbalances between the two states.

The deviation of the clear hole position caused by processing, aging, and thermal expansion can be regarded as short-lived, equal-width positive and negative pulse interference, while noise and other interference can be regarded as transient positive or negative pulse interference.

Although these signals are higher frequencies than ideal signals in the frequency domain, the simple use of frequency domain filters is not appropriate given that the spectrum of these signals still overlaps greatly with the ideal signal and that the signal frequency may vary greatly.

However, in the assumptions above, we believe that there is a clear difference in time domain characteristics between the error signal and the ideal signal, and we design an algorithm that can distinguish them and give a simulink model.

DOI: 10.54254/2755-2721/49/20241194

3. Signal processing procedure

From the assumptions above, it can be concluded that the ideal position signal should have approximately equal levels of time at high and low levels. A duty cycle correction (DCC) circuit is a widely used circuit that requires an accurate duty cycle clock, such as (DDR) DRAM, ADC, etc[7]. The clock duty cycle corrector calibrates the duty cycle of the clock to a value (typically 50%), so by drawing on the idea of clock duty cycle correction circuit design, we propose the following signal processing flow:

1. If the duration of the current high/low differs from the duration of the previous low/high by a small range, the signal is output as is.

2. Select a short time as the allowable change time, take the last high/low duration plus the allowable change time as the latest allowable time of the new edge, and the last high/low duration minus the allowable change time as the earliest allowable time of the new edge. If a new edge of the source signal appears prematurely, a new edge is made at the earliest allowable time, and if the new edge of the signal does not appear after the latest allowable time, a new edge is created at the latest allowable time.

Suppose an encoder has an output signal of 5us high and low levels, and due to manufacturing accuracy or interference, one of the channel high and low levels changes to 4us and 6us in one cycle. Setting tr to 0. 1us, we can compare the difference between using this method to perform a fourfold subdivision count of the processed signal and a fourfold subdivision count of the original signal.

$(\theta = \frac{2\pi}{4*\mathrm{PPR}})$

Table 1. Comparison of errors between the original signal and the processed signal.

Measured Position(θ)	True Position(θ) (original signal)	Error(θ) (original signal)	True Position(θ) (processed signal)	Error(θ) (processed signal)
1	1	0	1	0
2	2	0	2	0
3	2. 6	0. 4	2. 96	0. 04
4	4	0	4	0
5	5	0	5	0

It is not difficult to see that by limiting the time range of high and low levels, the error due to accidental changes in the signal can be reduced according to table 1. The part that implements this process is named OEC (Occasional Error Compensator).

However, a persistent imbalance between the durations of the two phases is also common in encoders, in which case the above processing method is likely to lose the ability to track the signal due to the long-term difference in the duration of the high and low levels, resulting in the complete disappearance of the role of the measurement position. Unfortunately, this phenomenon is quite common in square wave encoders, so compensating for it is necessary.

Some literature suggests a solution to the duty cycle imbalance, namely the analog DCC circuit mentioned above, which converts the duty cycle into a voltage quantity through a low-pass filter or integrator and establishes negative feedback to control it to 50%. However, these circuits are designed for high-speed circuits with frequencies above Ghz, in which case the clock signal is no longer an ideal square wave but a trapezoidal wave [8-10]. The frequency of the encoder output is mostly Mhz and below, and the rise and fall time is negligible. Therefore, by borrowing the design ideas of DCC circuits, SB (State Balancer) is redesigned to balance the gap between high and low levels.

3.1. State Balancer Design

An obvious result is that if the duty cycle of the square wave is not the ideal 50%, then there must be a duty cycle of less than 50% between the square wave and the negated square wave. If the duty cycle of the square wave is less than 50 percent, then its common-mode voltage must be less than the common-mode voltage after passing through the inverter, and vice versa, it must be greater than its common-mode voltage after passing through the inverter. Therefore, by comparing the common-mode voltage, a square wave signal with a duty cycle of less than 50% can be selected. We can complete the balance of

DOI: 10.54254/2755-2721/49/20241194

the high and low state by obtaining the time difference between the high and low levels of the signal and extending the high level by half of the time difference. However, because the frequency of the encoder changes, it is not suitable to directly obtain its high and low time for correction, so we design a negative feedback loop to compensate for the high level time as we can see in figure 1.

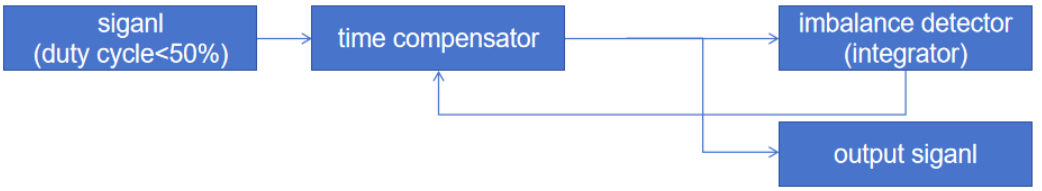

Figure 1. The overall structure of SB.

The time compensator begins integration at the end of the high level of the source signal, and if the integration reaches the control voltage of the time compensator, the compensation ends. The integrator integrates the output waveform, if the duty cycle of the output waveform deviates by 50%, it will cause the integration result to increase or decrease, and the integration result is negative as the control voltage of the time compensator, which can automatically lock the compensation time, and output a square wave with a duty cycle of 50%.

We built a simulation model of SB using simulink and created a signal source with a duty cycle of 40% and a continuous frequency change to simulate the encoder to measure the state during uniform acceleration motion. Its frequency changes continuously from 10Hz to 110Hz within five seconds. After the transient time after about 0. 9s, the duty cycle is adjusted to 50%. Extracting the common-mode voltage of the output signal through a low-pass filter shows that a duty cycle of close to 50% is maintained over the entire frequency range.

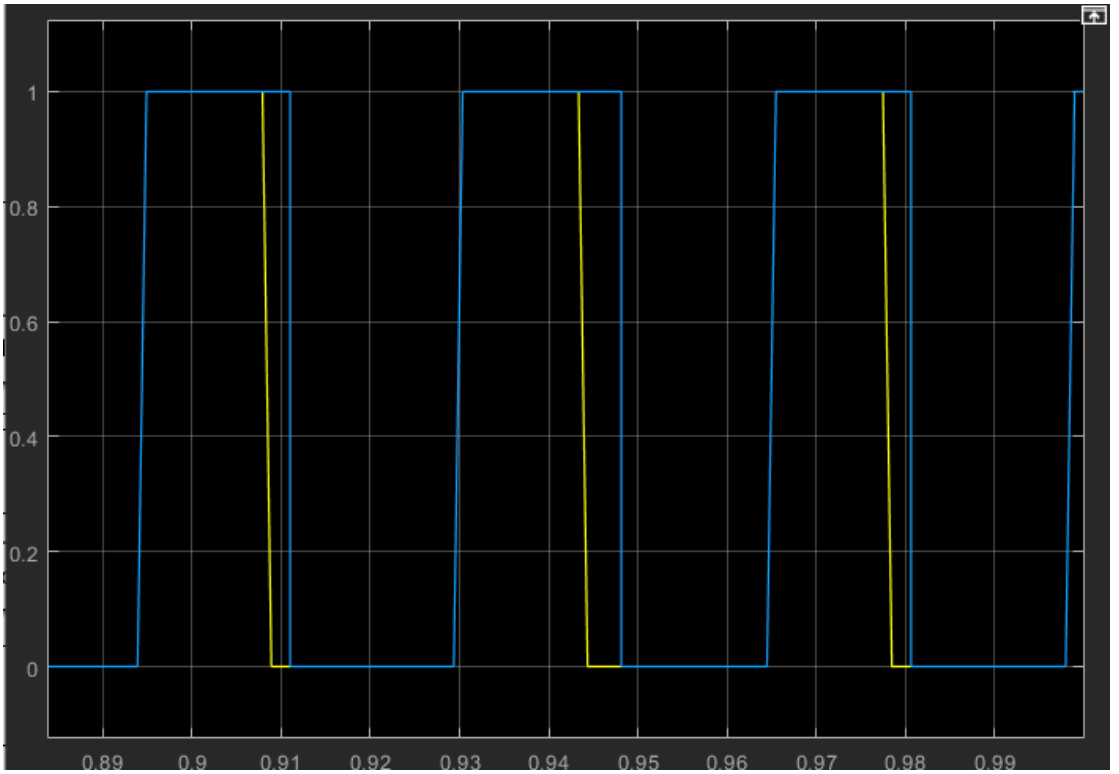

Figure 2. (30hz) output signal(blue) vs. source signal(yellow).

DOI: 10.54254/2755-2721/49/20241194

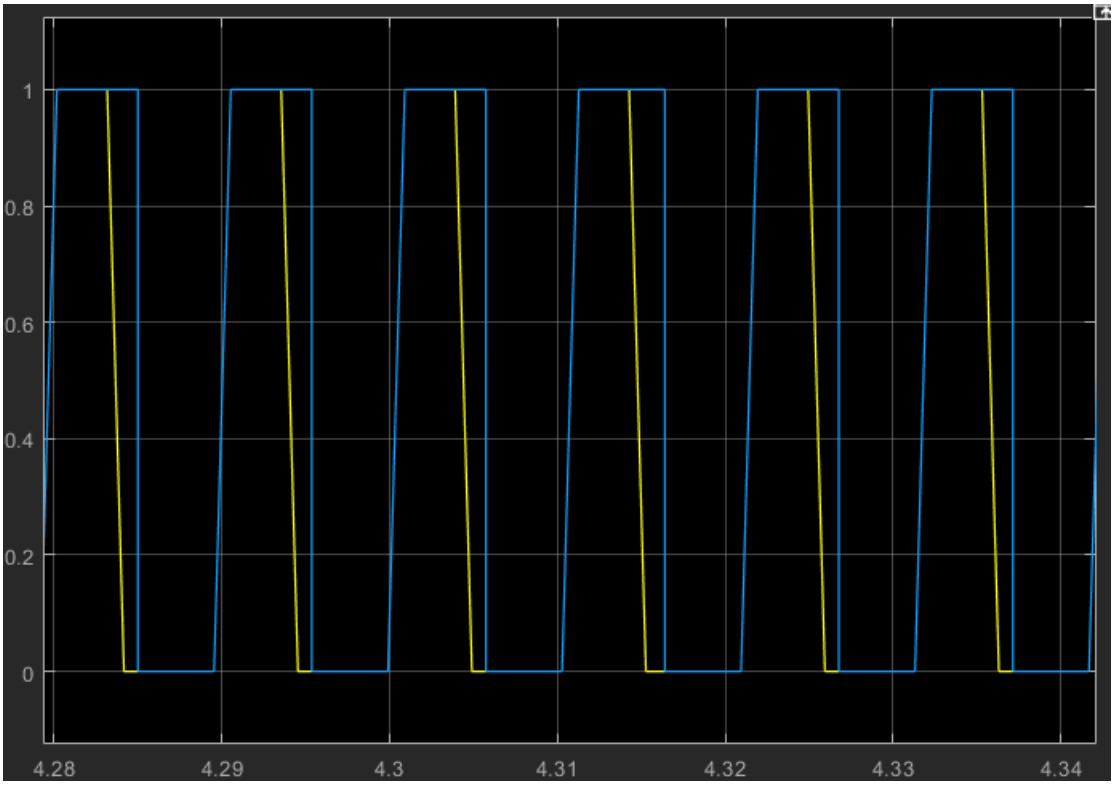

Figure 3. (105hz) output signal(blue) vs. source signal(yellow).

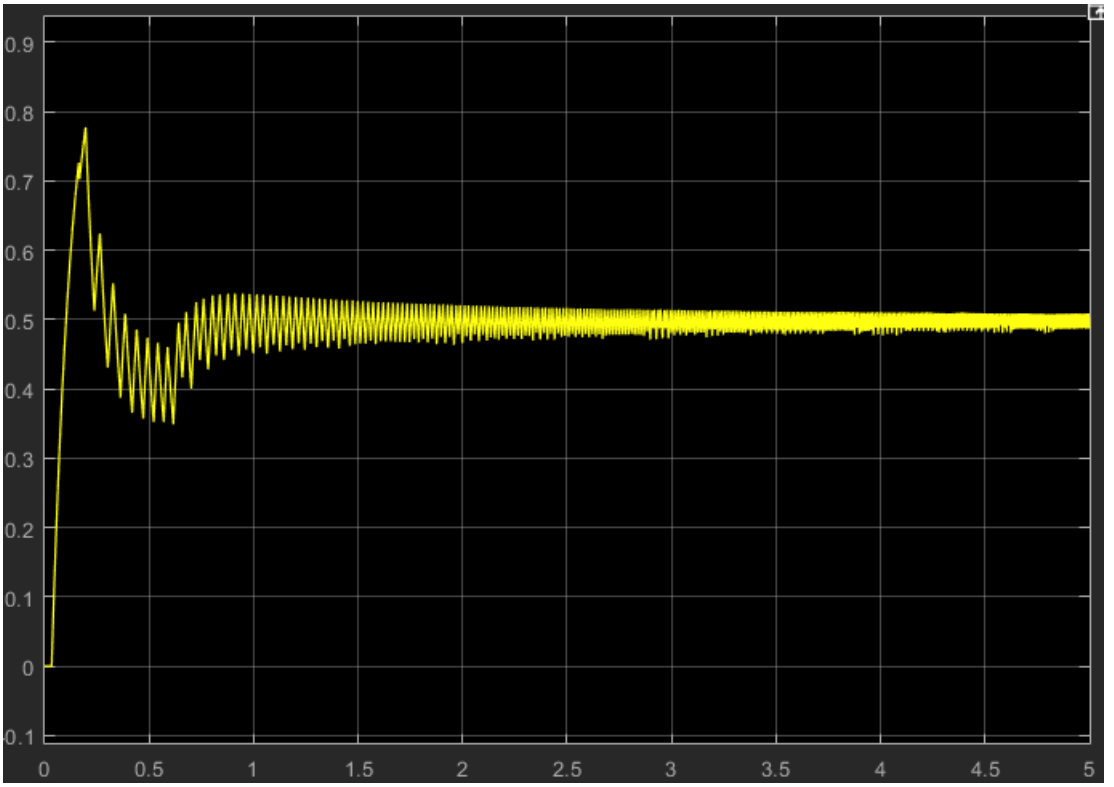

Figure 4. common-mode voltage of the output signal.

DOI: 10.54254/2755-2721/49/20241194

In SB, the response speed can be improved by increasing the integral gain of the detection link and reducing the integral gain of the compensation link, but the fluctuation of the duty cycle will be increased. On the contrary, reducing the gain of the detection link, increasing the gain of the compensation link, and adding the dead zone of the compensation link can obtain a more accurate duty cycle, but the adjustment time is longer as shown in figure 2 to 4.

As mentioned above, the state balancer cannot properly compensate for the pulse interference in time due to the low-pass nature of the detection process.

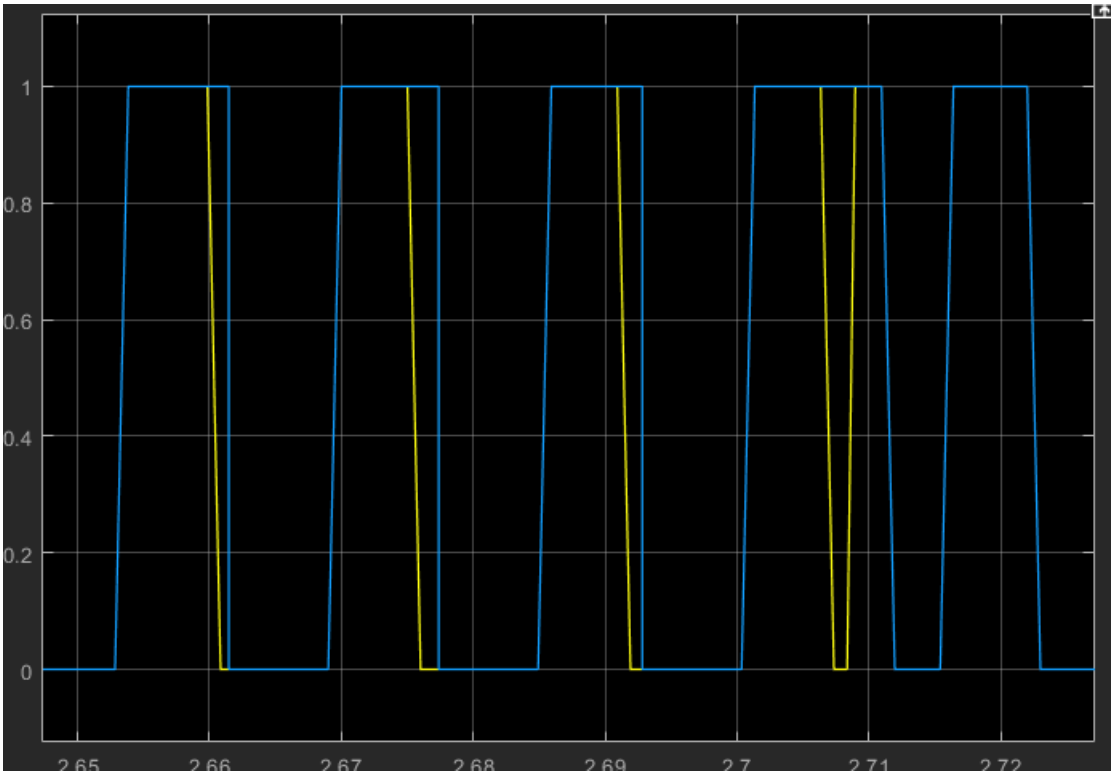

Figure 5. Output signal under rectangle pulse interference.

Accordig to figure 5, this method of signal processing has the advantages of not requiring accurate resistor and capacitor values, having a wide frequency processing range, and not requiring an external high-speed clock, so it can be achieved using only op amps, flip-flops, and RC circuitry.

3.2. Occasional Error Compensator Design

Based on the signal processing flow given above, we can implement it with a dual-channel coupled integral time limiter, which works as follows:

A channel accesses the square wave signal, the signal is recorded as SA, and the B channel accesses the square wave signal after the inverse, and the signal is recorded as SB. Both A and B are connected to an integral circuit that is charged and discharged at a constant rate.

The triangular wave timing signal output by the A integrator is recorded as UA, and the triangle wave signal of B channel is recorded as UB.

Both channels use a trigger to enable after their first rising edge and remain enabled thereafter until the end of processing.

Set a small voltage Uset.

Record the trigger function as FA, and the FA trigger condition is ((SA posedge and UA<Uset) or UA<-Uset)

Similarly, FB trigger conditions are ((SB posedge and UB<Uset) or UA<-Uset)

DOI: 10.54254/2755-2721/49/20241194

After enabling, when FA is triggered, the UA voltage is quickly cleared, UB enters the discharge state, the UA enters the charging state, and the output signal changes to high. In the same way, when FB is triggered, the UB enters the charging state, the UA enters the discharge state, and the output signal changes to low.

The time sustained by the previous high/low is compared to the next low/high by timing a constant-rate charged triangular wave signal, and if the time difference is within a tolerable range, the edge of the original signal is reproduced. If the time edge occurs before the time has arrived, or if the allowable time has passed and the edge does not appear, a new edge is made with its allowable lower/upper time. The time error caused by this occasional interference can only have a very limited effect on the signal time, and does not interfere with the signal output under normal conditions.

Simulink models this process and tests its ability to eliminate square wave interference as shown in figure 6.

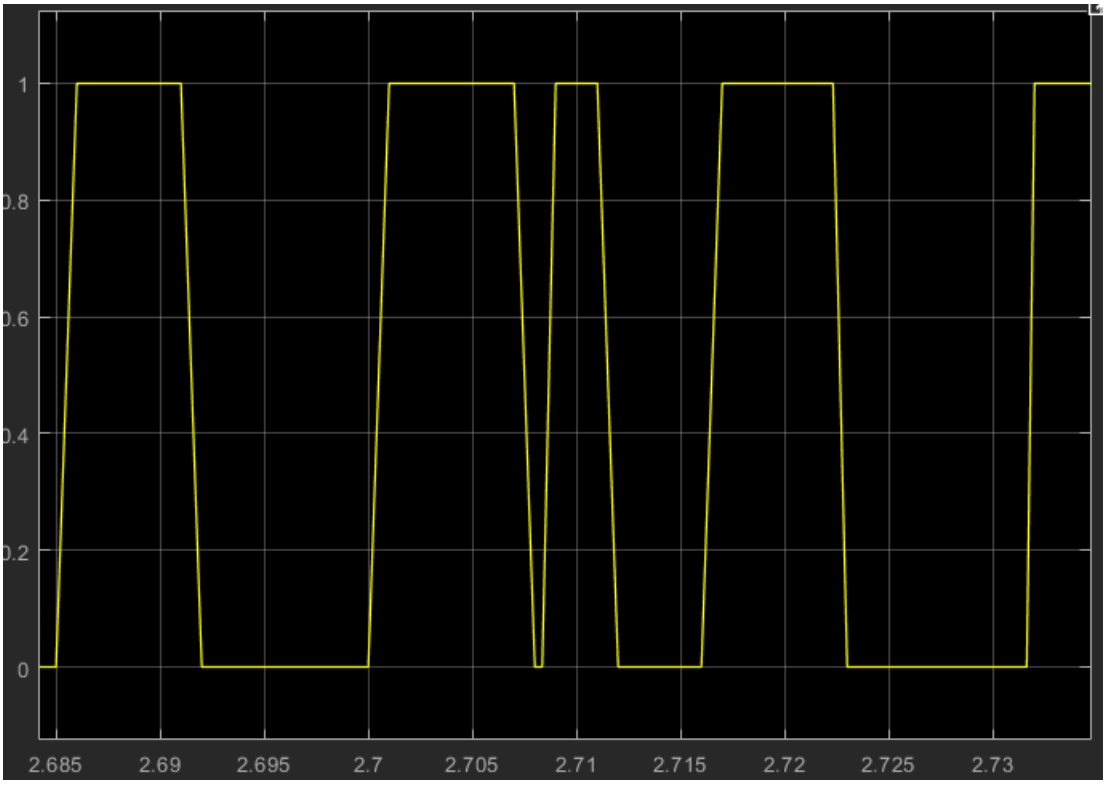

Figure 6. Source signal with rectangle pulse interference and state imbalance.

DOI: 10.54254/2755-2721/49/20241194

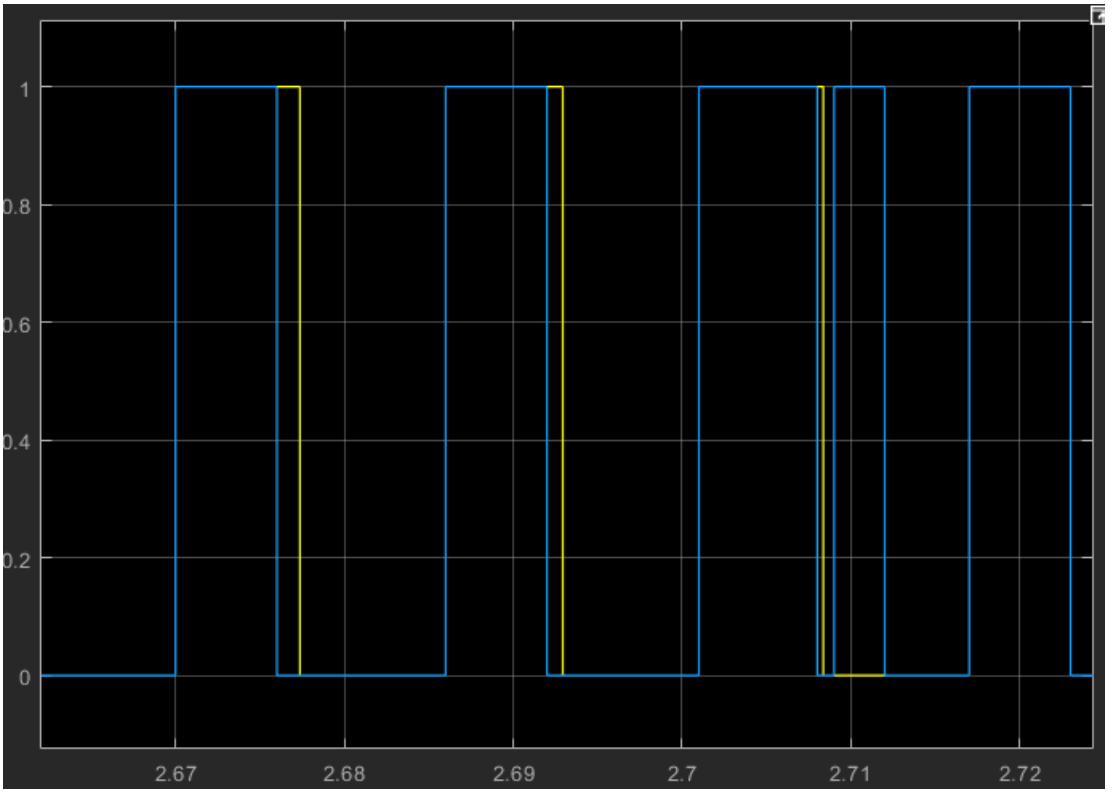

Figure 7. Signals processed by SB and OEC(blue) vs source signal(blue).

It is not difficult to see that the signal processed by SB and OEC can better overcome the state imbalance and occasional interference of the encoder signal, greatly improve the signal quality, reducing error to 5% of PPR and provide a guarantee for reliable and accurate higher subdivision according to figure 7.

This signal processing method has the following advantages: since the signal can only be charged until it is discharged until near the zero point, the zero drift of the integration circuit is suppressed; Two-way integrators do not require precise resistor and capacitance values; No external high-speed clock is required, so it can be achieved using only op amps, flip-flops, and RC circuitry.

4. Conclusion

By studying the characteristics of the ideal encoder signal, we pointed out that the ideal position signal output by the encoder can be approximately regarded as a square wave signal with a duty cycle of 50%, and the effectiveness of this signal processing method is verified by designing SB and OEC for compensation based on the difference in the time domain characteristics of the ideal position signal and the interference signal, and simulating it by using simulink. This processing method has the advantage that it does not require precise resistor and capacitance values, does not require external high-speed clock access, and can be implemented by basic components. It is worth noting that when the encoder is running at very low speeds, the assumption that the output square wave duty cycle should be 50% no longer holds, so this processing method will fail. There are still many shortcomings in the current work: encoders operating at low speeds cannot be measured, the duty cycle cannot be controlled to 50% absolute due to limited high-frequency rejection capabilities, accuracy and speed cannot be guaranteed at the same time, and there is a two-step long transient time before establishing stable error compensation.

DOI: 10.54254/2755-2721/49/20241194

References

[1] Liu, R., & Liu, R. (2016). Signal acquisition technology for photoelectric encoder based on FPGA. Optik, 127, 9891-9895.k 127 : 9891-9895.

[2] Jiang, Jiaqi & Dai, Jiahai & Yang, Shang & Chang, Yunchun. (2022). A 22-bit image encoder with optoelectronic integrated chip. Optics Communications. 512. 128022. .

[3] H. Rieder, M. Schwaiger, RSF. Elektronik, GmbH. Tarsdorf and Austria,(1991) "Method of Electronically Correcting Position Errors in an Incremental Measuring System and Measuting System for Carrying out the Method", United States Patent: 5021650, June .

[4] T. Emura, L. Wang and A. Arakawa, (1993) "A High-Resolution Interpolator for Incremental Encoders By Two-Phase Type PLL Method", Proceedings of the International Conference on Industrial Electronics Control and Instrumentation, vol. 3, pp. 1540-1545, .

[5] Chi Feng Lan, Development of Electronic Subdivision with FPGA and Its Application in Grating Interferometer, Taipei, Taiwan:National Taiwan University, 2011.

[6] M.-S. Wang, Ying-Shieh Kung, Yi-Ming Tu and Tsung-Ting Lin, 2009 "Novel interpolation method for quadrature encoder square signals," In IEEE International Symposium on Industrial Electronics, Seoul, Korea (South),, pp. 333-338

[7] Y. -J. Min et al. , 2012 "A 0. 31–1 GHz Fast-Corrected Duty-Cycle Corrector With Successive Approximation Register for DDR DRAM Applications," In IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 20, no. 8, pp. 1524-1528,

[8] Sim, J., Park, H., Kwon, Y., Kim, S.W., & Kim, C. 2021 A 1-3.2 GHz 0.6 mW/GHz Duty-Cycle-Corrector Using Bangbang Duty-Cyle-Detector. In IEEE International Symposium on Circuits and Systems (ISCAS), 1-4.

[9] C. Y. Lin and H. S. Hsu, "Design of A 0. 8GHz-3GHz Duty-Cycle Corrector With a 20%-80% Input Duty Cycle," 2019,In IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW), Yilan, Taiwan, pp. 1-2.

[10] Prakash S.R. and S. S. Hiremath, "Dual loop clock duty cycle corrector for high speed serial interface," 2017 In International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bengaluru, India, pp. 935-939,