Video over SPI
(VoSPI)
Implementaion
Specification

Document Number: 102-PS245-43
Date: 22 Jan 2013

Weilming Sieh, FLIR

Paul Fagerburg, Syncroness

Reviewed by:

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. DOCUMENT

1.1 Revision History

Rev	Date	Comments
001	Aug 11, 2011	Creation
002	Nov 14, 2011	First Draft Release
003	Dec 21, 2011	Fixed Header and Footer
004	Feb 3, 2012	Changed Q to T in Video packet. Changed EOF to look more like SOF
005	22 Jan 2013	Made the document Lepton-specific. Added details for the SPI bus. Clarified discard packets.
006	Mar, 1, 2013	Re-titled document and gave document a new CC number. Removed language discard packet not containing a valid CRC. The non-valid CRC is a artifact of a bug in the Lepton emulation platform.
007	Sep 9, 2013	Updated CRC Polynomial to use CCITT16 Kermit

Table 1 Revision History

List of open / unresolved items in the current draft:

Item	Comments

Table 2 Open Issues

2. INTRODUCTION TO FLIR VOSPI

FLIR's Video over SPI (VoSPI) Protocol is designed to encapsulate video in a format that allows transmission over a SPI interface while requiring minimal software or hardware respectively. There are two types of systems for which this protocol is optimized:

- A software system will typically have an off-the-shelf commercial processor attached to a display. Systems in this class typically use off-the-shelf commercial processors and display components. VoSPI does not require the processor to "touch" the pixel data, nor to spend time searching through the data for VSYNC, HSYNC and padding bits, which greatly reduces the CPU load to receive the video and send it to the display. If the processor has a DMA controller, VoSPI requires minimal CPU overhead.

- A hardware system will typically use custom logic to receive and render the video. VoSPI requires very little logic - only a programmable frame time counter/divider, and a 4-bit comparator. The protocol avoids embedded timing signals and padding bits to prevent fractional bit-time resolution errors accumulating into video-tearing artifacts across multiple frames.

VoSPI supports multiple levels of functionality, including advanced features that are not within the scope of this document.

This document only covers the Lepton camera's implementation of VoSPI.

2.1 VoSPI Terminology

VoSPI uses the same master/slave terminology as the underlying physical layer, SPI. A device that generates a video image (i.e. a camera) is designated as a VoSPI slave. A device that receives a VoSPI stream is designated as a VoSPI master.

2.2 Physical Implementation

VoSPI uses a modified Serial Peripheral Interface (SPI) to carry the data. The SCK (Serial Clock), /CS (Chip Select, active low), and MISO (Master In/Slave Out) signals are present, but VoSPI does not use the MOSI (Master Out/Slave In) signal. Implementations are generally restricted to a single master and a single slave, due to the strength of the line drivers in the hardware.

The Lepton operates as a slave port only. The host controller must operate as an SPI master to generate the SCK (Serial Clock) signal.

The Lepton uses SPI Mode 3 (CPOL=1, CPHA=1). SCK is high when idle. Data is set up by the Lepton on the falling edge of SCK, and should be sampled by the host controller on the rising edge of SCK.

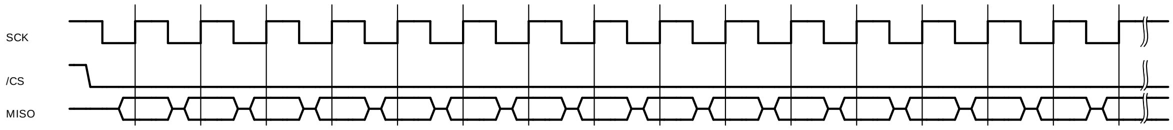

Figure 1 SPI Mode 3

Data is transferred most-significant byte first and in big-endian order; the value 0x8C08 will appear on the MISO line as shown below:

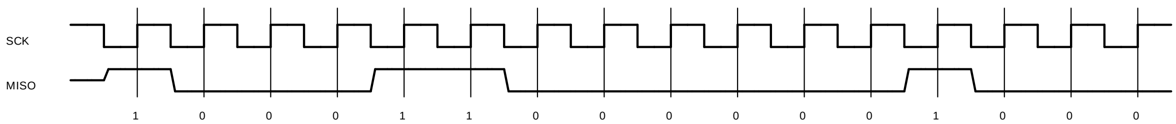

Figure 2 SPI Bit Order

3. VOSPI DEFINITIONS

VoSPI is built on a collection of object types as defined hierarchically below.

3.1 VoSPI Packet

The VoSPI is based on a single standardized VoSPI Packet, which is defined as a 164 byte packet. When transferring a packet, the VoSPI master must always transfer the entire packet.

Packets can contain pixel information (Video Packets) or stream and format information (Header Packets). Header packets are outside the scope of this specification.

3.2 VoSPI Frame

A VoSPI Frame is defined as a continuous sequence of VoSPI packets. The VoSPI master can determine the Start of Frame from the transmission of Line 0, and the height (H) from the highest line number transmitted just before Line 0.

3.3 VoSPI Stream

A VoSPI Stream is defined as a continuous sequence of VoSPI Frames.

4. PACKET TYPE DEFINITIONS

The Lepton implementation of VoSPI is very simple, consisting only of video packets and discard packets. Other packet types are outside the scope of this document.

ID	CRC	PAYLOAD
4 Bytes		160 Bytes

Figure 3: Generic VoSPI Packet

4.1 CRC Calculation

The packet header contains a 16-bit CRC value, using the following polynomial:

$$x^{16} + x^{12} + x^{5} + x^{0}$$

The CRC is calculated over the entire packet, including the ID and the CRC fields. However, the four most-significant bits of the ID and all sixteen bits of the CRC are set to zero as part of the calculation of the CRC. Note that in hardware display applications, it will be necessary to save (register) the CRC at the head of the packet for comparison with the re-generated packet CRC.

4.2 Video Packet

ID	CRC	PAYLOAD
TNNN	CRC	Video pixels for Line NNN

Figure 4: Video Packet

The ID field consists of 1 reserved bit, a 3-bit time field, and a 12-bit line number. In this example, the T field is 7, and the NNN field is 4:

0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0
	T			NNN											

The T field is the 24 LSB's of the current Frame Time, shipped 3 bits at a time, LSB first, over 8 consecutive packets. The first video packet of a frame always restarts from the LSB of the current frame time. If the number of lines (packets) per frame is not divisible by 8, the remaining bits are NOT shipped in video packets of the next frame. Packet 0 gets bits 2..0, packet 1 gets 5..3, packet 2 gets 8..6, and so on until packet 7 gets bits 23..21. Packet 8 gets bits 2..0 again, and so on.

The payload contains 80 pixels (160 bytes) of video data.

4.3 Discard Packet

ID	CRC	PAYLOAD
xFFx	CRC	DISCARD DATA

Figure 5: Discard Packet

The “x” in the ID field signifies a “don’t care” condition. If bits 4 through 11 are set, then the packet can be discarded without examining bits 0-3 and 12-15. Because VoSPI-enabled cameras do not have a vertical resolution even close to 4080 lines (0xFF0), there is no possibility of the ID field in a discard packet being mistaken for a line of video.

Discard packets may contain additional information about the state of the VoSPI slave and the video stream, but the VoSPI master is not required to use any of it to properly decode the stream.

5. LEPTON VOSPI IMPLEMENTATION DETAILS

Lepton implements the minimum level of the VoSPI protocol. Each packet contains a single line of video, 80 pixels of 16 bits each. With the ID and CRC fields, a Lepton VoSPI packet is 164 bytes in length. Therefore, a VoSPI master must always transfer 164 bytes of data in a single SPI transaction. Refer to Figure 7 below.

A single frame contains 60 lines. At the beginning of SPI video transmission, Lepton will send discard packets until it has a new frame from its imaging pipeline. It will then send the 60 lines of video data in order.

ID	CRC	PAYLOAD
xFFx	CRC	DISCARD DATA

.
.
.

ID	CRC	PAYLOAD
xFFx	CRC	DISCARD DATA
T000	CRC	Line 000: P0 | P1 | P2 | P3 | P4 | P5 | ... … ... | 78 | 79
T001	CRC	Line 001: P0 | P1 | P2 | P3 | P4 | P5 | ... … ... | 78 | 79
T002	CRC	Line 002: P0 | P1 | P2 | P3 | P4 | P5 | ... … ... | 78 | 79
T003	CRC	Line 003: P0 | P1 | P2 | P3 | P4 | P5 | ... … ... | 78 | 79
T004	CRC	Line 004: P0 | P1 | P2 | P3 | P4 | P5 | ... … ... | 78 | 79

.
.
.

ID	CRC	PAYLOAD
T058	CRC	Line 058: P0 | P1 | P2 | P3 | P4 | P5 | ... … ... | 78 | 79
T059	CRC	Line 059: P0 | P1 | P2 | P3 | P4 | P5 | ... … ... | 78 | 79

Figure 6: Lepton Example SPI Video

If the VoSPI master does not start transferring video packets within 2-4 line times of VSYNC, Lepton will revert to its initial state, transmitting discard packets, in an attempt to sync.

The chip select signal ends the transfer of the current packet. The VoSPI master should leave chip select de-asserted for a minimum time of the equivalent of 64 SCKs, even though SCK is not being driven to the VoSPI slave; i.e. if SCK is 4 MHz, then /CS should be de-asserted for at least 64/4MHz = 16µs. Refer to Figure 8 below.

If the VoSPI master does not transfer a complete packet, the VoSPI slave may return a misaligned packet for the next transfer, rather than a discard packet. In this case, the master must de-assert /CS for at least 1 second to allow the VoSPI slave logic to time out and reset.

The Lepton has an internal frame rate of 27Hz, but the frame data is updated only on every third frame, giving an effective frame rate of 9Hz for ITAR compliance. The VoSPI master should transfer all 60 lines of video data before the next 27Hz interval (37ms). Figure 9 below shows the transfer of an entire frame of video at a lower level of detail. The 27Hz frame timing is shown with the transfer of individual lines of video data. Each line of video is shown as a number, with “D” signifying a discard packet.

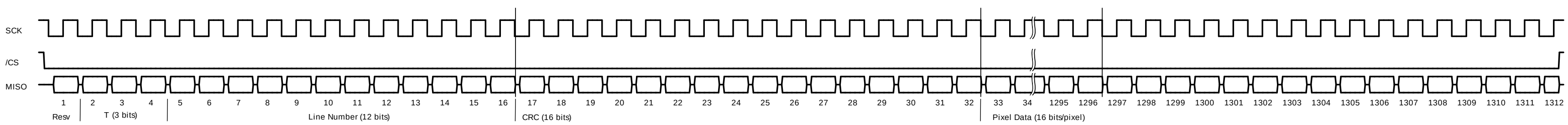

Figure 7 Complete VoSPI Packet Transfer

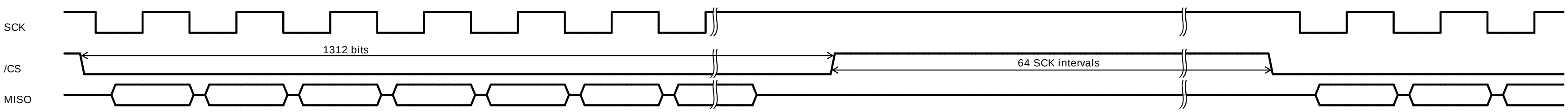

Figure 8 Multiple VoSPI Packets with Inter-Packet Time

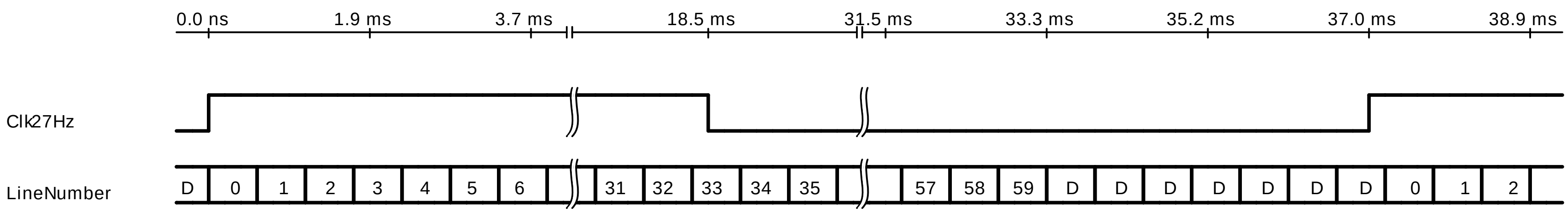

Figure 9 Transferring an Entire Frame, With Discard Packets

5.1 **Test Pattern Data**

The Lepton camera can output a test pattern where each pixel takes the value of line*100+column. This test pattern is very easy to spot on a logic analyzer when examining the MISO line from the Lepton.

The following data was captured from a Lepton emulation platform (ML605 FPGA Evaluation board) using custom logic.

```
3000 8c08 0000 0001 0002 0003 0004 0005 0006 0007 0008 0009 000a 000b
000c 000d 000e 000f 0010 0011 0012 0013 0014 0015 0016 0017 0018 0019
001a 001b 001c 001d 001e 001f 0020 0021 0022 0023 0024 0025 0026 0027
0028 0029 002a 002b 002c 002d 002e 002f 0030 0031 0032 0033 0034 0035
0036 0037 0038 0039 003a 003b 003c 003d 003e 003f 0040 0041 0042 0043
0044 0045 0046 0047 0048 0049 004a 004b 004c 004d 004e 004f
```

Figure 10 Raw Data Capture of Test Pattern Line 0

The first word, 0x3000, signifies a T value of 3 and a line number of 0. The next word, 0x8c08, is the CRC of the entire packet if the T and the CRC are both set to zero before computing the CRC. The next 80 words are the values 0 through 79, according to the test pattern formula line*100+column.

The pattern continues to the point where the VoSPI slave outputs the last line of data.

```
003b 48d0 170c 170d 170e 170f 1710 1711 1712 1713 1714 1715 1716 1717
1718 1719 171a 171b 171c 171d 171e 171f 1720 1721 1722 1723 1724 1725
1726 1727 1728 1729 172a 172b 172c 172d 172e 172f 1730 1731 1732 1733
1734 1735 1736 1737 1738 1739 173a 173b 173c 173d 173e 173f 1740 1741
1742 1743 1744 1745 1746 1747 1748 1749 174a 174b 174c 174d 174e 174f
1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 175a 175b
```

Figure 11 Raw Data Capture of Test Pattern Line 59

The first word, 0x003b, has a T value of 0 and a line number of 59. The next word, 0x48d0, is the CRC, as describes earlier. The next word, 0x170c or 5900 decimal, matches the formula for computing the pixel values. This pattern continues until the value 0x175b (5979 decimal), which is line 59, column 79, the last pixel in the test image.

After outputting line 59, the VoSPI slave will output discard frames until the next 27Hz interval, at which point it will begin outputting line 0 of the test pattern again.

```
1fff ffff dcd0 dcad dcd2 dcad dcd4 dcad dcd6 dcad 0000 0000 0000 0000
0000 2ebc 3b00 1042 085e 0100 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
```

Figure 12 Raw Data Capture of a Discard Packet