CMOS ACTIVE PIXEL SENSOR TECHNOLOGY
- AN INTRODUCTION -

Dr. Eric R. Fossum
Imaging Systems

Optics

Timing & Driver Circuits

Detector Pixel Array

Analogue Signal Processor
Analogue to Digital Converter
Digital Signal Processor

ASP
ADC
DSP
- A photon is a “bullet” of light (quantum view of light waves)
- Energy is related to wavelength: \( E = \frac{hc}{\lambda} \)

Your eye is sensitive from violet to dark red
- 450 nm - 700 nm

Silicon is sensitive from violet to the “near” infrared (not thermal)
- 400 nm - 1050 nm

Silicon is also sensitive to higher photon energies (e.g. x-rays)
Sources of Photons

- Three main sources
  - Solar
  - Incandescent lamp
  - Fluorescent lamp
- Typically characterized by ‘color temperature’
  - Temperature of blackbody with roughly same color distribution
  - Higher temp means more blue, less red -> looks white
SPECTRAL RESPONSES

- Human Eye
- Silicon

Wavelength (nm):
- UV
- NIR

Relative response:
- 0
- 0.1
- 0.2
- 0.3
- 0.4
- 0.5
- 0.6
- 0.7
- 0.8
- 0.9
- 1

© ER FOSSUM
Optics

- Optics (lenses, mirrors) COLLECT and FOCUS light
- Optics are measured by focal length (zoom-effect)
- and by “speed” or F-number (ratio of focal length to diameter)
  - Light on sensor is $1/4F^2 \times$ light on scene
    - Higher F-number means less light on image sensor
    - Lower F-number mean more light
  - Iris or diaphragm adjusts F-number by changing effective lens diameter
  - Typical video cameras are F/2 (so ‘faceplate’ illuminance is 1/16 scene)
- Standard lens sizes are measured in inches (sort of lens diameter)
  - e.g. 2/3”, 1/2”, 1/3”, 1/4” etc.
  - Smaller lens sizes weigh less and cost less - important for portables
  - Smaller sizes require smaller image sensors and smaller pixel sizes (for same # of pixels) and are less sensitive to light
  - Actual image sensor diagonal is much less than lens size
  - For example, 1/3” format (8.5 mm) needs 6.1 mm diagonal!
A pixel is a “picture element”
More pixels means more RESOLUTION
More pixels means more DATA (could be a bad thing)
Modern pixel sizes are between 4 and 10 microns
Pixels are square if designed for a computer display
Pixels have a rectangular shape for TV applications

<table>
<thead>
<tr>
<th>FORMAT</th>
<th>Columns</th>
<th>Rows</th>
<th>Total # Pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIF</td>
<td>352</td>
<td>288</td>
<td>100k</td>
</tr>
<tr>
<td>VGA/TV</td>
<td>640</td>
<td>480</td>
<td>300k</td>
</tr>
<tr>
<td>Megapixel</td>
<td>1280</td>
<td>1024</td>
<td>1.3M</td>
</tr>
<tr>
<td>HDTV</td>
<td>1920</td>
<td>1120</td>
<td>2.0M</td>
</tr>
</tbody>
</table>
• Each pixel gets covered by a colored filter
  – We use red, green, blue (RGB) CFA - best match for RGB displays in “Bayer” pattern
  – Could use complementary colors (cyan, yellow, magenta)
Fill Factor

- A pixel is divided into a sensing portion and a readout portion.
- Fill factor is the ratio of sensing area to total area and is typically about 20% for CCDs and CMOS APS.
• Microlenses funnel light away from non-sensitive portions of pixel
• Photon is absorbed by silicon and converted to an electron-hole pair. Typically the electrons (e-) are collected and holes discarded.
• Quantum efficiency (QE) is ratio of collected electrons to incident photons over whole pixel* and is always less than unity for visible light, and is wavelength dependent.
• Depends on:
  – fill factor
  – microlenses
  – design
  – vertical photosensor structure
  – fundamental physics of silicon

* some people just use the defined sensitive area to get a higher value but we consider this cheating. One always needs to check how it is defined.
APS has quantum efficiency comparable to CCDs
Each pixel has its own output amplifier
- Pixels are X-Y addressed
- Key is low noise readout circuit
- Best of CCD detection/readout and CMOS integration
The CONVERSION GAIN determines the amount of volts per electron for the pixels.  
Typical value range is 1-10 μV/e -  
Higher value is good but limits total signal that can be handled  
- e.g. at 10 μV/e -, 1 volt max = 100,000 electrons max  
Pixel amplifier should introduce minimum noise  
Noise is measured in equivalent number of electrons  
- e.g. 200 μV rms @ 10 μV/e - = 20 e - rms noise
The ANALOG-TO-DIGITAL CONVERTER (ADC) converts the voltage from the pixels to a digital word.

The ADC may have an ANALOG SIGNAL PROCESSOR (ASP) to reduce noise and provide additional gain before conversion.

ADC has a reference voltage, e.g. 1 Volt, so that digital word is scaled against the reference voltage. A lower reference voltage is similar to adding gain.

On-chip ADCs trade power for resolution (number of bits)

8 bits is enough for teleconferencing. 10-12 bits is desired for digital still cameras.
Photons to Bits

1/4F²

Scene to Sensor

Funnels Light

Selects Color

Includes Fill-Factor (no cheating)

Electrons to Volts

Amplification and Noise Suppression

Volts to Bits

OPTICS ➔ MICRO-LENS ➔ COLOR FILTER ARRAY (CFA) ➔ QUANTUM EFFICIENCY (QE) ➔ CONVERSION GAIN (uV/e⁻) ➔ ASP ➔ ADC

© ER FOSSUM
Sensor produces 8b (or more) for each pixel
System eventually wants 24b (RGB) for each pixel
  – Color interpolation required
  – Color balancing (“white balance”) required
Additional processing used for aperture correction, etc.
Communication is usually bandwidth limited
  – Compression often required
Interfaces need to have data format and output according to interface specification
  – Interface controller/data formatter required e.g., USB, NTSC
Color Interpolation

- Goal is to get best approximation for RGB at each pixel site
- Many possible approaches, e.g.:

  • Have blue, need green & red
    - G = average of 4 neighboring greens
    - R = average of 4 neighboring reds

  • Have green, need blue & red
    - B = average of 2 neighboring blues
    - R = average of 2 neighboring reds

  • Have red, need green & blue
    - G = average of 4 neighboring greens
    - B = average of 4 neighboring blues
Color Correction

- Need to correct colors
  - relative intensities (white balance)
  - remove mixtures (off-diagonal elements)

\[
\begin{bmatrix}
R' \\
G' \\
B'
\end{bmatrix} =
\begin{bmatrix}
r_r & r_g & r_b \\
gr & g_g & g_b \\
b_r & b_g & b_b
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

No perfect solution due to filter overlaps and eye response
Best-fit by weights to standard color chart
Interpolation and color correction leads to ‘blurring’ of image since averaging over a neighborhood.

Want to restore apparent sharpness to image.

Apply simple image sharpening algorithm to ‘green’ image.

Apply green correction to red and blue.
• Communications typically limited in bandwidth (bits/sec)
• USB, Firewire, POTS, Internet….
• Many different compression techniques
  – JPEG - still pictures
  – MPEG - moving pictures (e.g. video)
  – wavelet
  – fractal
  – vector quantization
• Few lossless approaches - image quality ALWAYS degrades -- but is it acceptable?
• Poor sensor plus poor compression = very poor image
• How design and partition system? Many options.
• Jury is out on ASIC vs. PC-based processing.
Imaging System-on-a-Chip

**PROS**
- Possibly lower cost
  - single package
  - single-pass testing
- Smaller footprint / form factor
- Lower total power dissipation
- Easier system assembly
- Increased reliability

**CONS**
- Possibly higher cost
  - Color filters and microlenses deposited on digital logic
  - Use of larger than optimal design rules
  - Lower yield due to large die size
- Noise corruption due to on-chip digital clocking
- Power dissipation leads to increase in dark current level
• 1 Gb/s ~ 335 TeraBytes/month
• Total Worldwide Internet Traffic 2000*
  – 20,000 TeraBytes/month, or 60 Gb/s average load
  – increasing 100%/yr, and about 10% total capacity
• A single CIF resolution image sensor PC video camera with 3:1 compression produces about 10 Mb/s.
  – 6,000 households with live cameras are equivalent to the total worldwide internet traffic
• A 1.3 Mpixel sensor at 1000 fps with 10b output produces about 13 Gb/s
  – 5 High speed high resolution cameras are equivalent to total worldwide internet traffic

© ER FOSSUM