Multi-Aperture Imaging Devices

Keith Fife, Abbas El Gamal and H.-S. Philip Wong

Department of Electrical Engineering
Stanford University
Multi-Aperture System

- Scene focused via objective lens above detector plane
- Re-imaged via local optics onto disjoint arrays
- Arrays have overlapping fields of view
- Image is formed using digital signal processing
Traditional vs Multi-Aperture

Traditional optical configuration

Multi-aperture optical configuration
Why Multi-Aperture Imaging?

- Capture depth information
- Reduce requirements of objective lens (cheaper optics)
- Achieve better color separation (less crosstalk)
- Redundant data allows for manufacturing defect correction
- Facilitate new circuit design architectures
- Benefit from pixel scaling
Why Image Sensors in CMOS?

- Enable integration of analog and digital circuits along with the sensor
  - Camera on a Chip
- Leverage high-volume, mainstream process
  - Potentially lower cost
- Create new applications
  - High Speed
  - High Dynamic Range
  - Low power
  - Region of Interest and random access

Our goal is to use CMOS technology to create *multiple* cameras on a chip.
Multi-Aperture Architecture

- The sensor contains an $m \times n$ array of $k \times k$ pixel groups
Array Structure

- Single aperture array with local readout
- Architecture enables global exposure
Capture and Readout Sequence

- **Frame timing**

  - Reset
  - Integration
  - Transfer
  - Readout
  - Vblank
  - Transfer

  \[ T_{int}, T_{out}, T_{frame} \]

- **Row timing**

  - V
  - H
  - RT
  - TX
  - RS

  \[ S1, S2 \]
MOS Structure for Typical Process

- Silicided regions: S/D and Poly
- STI, Nitrided spacers, SiON for BLC
Typical 4T Pixel

- Diode is fully depleted of electrons during readout
- Surface of diode is pinned at substrate potential
- Charge to voltage conversion occurs at each pixel
- Floating diffusion (FD) is doubled sampled
Multi-Aperture Pixels

- CCD channels and channel stops created with CIS implants
- Polysilicon used in active array with no silicide
- Charge confined under electrodes and shifted to output
- No metal layers are needed in the imaging area
Array Layout

- Each array is a 16 $\times$ 16 Full Frame Transfer CCD
2D and 3D Image Extraction

- Depth information is obtained from the disparity between apertures.
- Object movement translates to lateral displacement between corresponding points imaged by disjoint arrays.
- Solving the correspondence problem is eased by using several local apertures.
- The 2D image is formed by solving for the local correspondence and integrating the result across the sensor.
Virtual Aperture Views

Chief rays for a pair of apertures

Left virtual objective aperture

Right virtual objective aperture

Virtual apertures for stereo view
Concept of Depth Extraction

Objects at varied distance

Objects captured with several apertures

Distances calculated from edge displacements
**Depth Calculations**

- By the geometry of the local optics and focal plane,

\[
\frac{C}{L} = \frac{D_0}{\Delta}
\]

- Using the lens law for \( A \) as a function of \( B \) and making the substitution

\[
B = E - C = B_0 + C_0 - C,
\]

\[
A = \left( \frac{1}{f} - \frac{1}{B} \right)^{-1} = \left( \frac{1}{f} - \frac{1}{B_0 + C_0 - C} \right)^{-1}
\]

- Solving for \( A \) in terms of \( \Delta \) with \( M = B/A \) and \( N = D/C \) gives the depth equation,

\[
A = \left[ \frac{1}{f} - \frac{1}{(M_0 + 1)f + D_0/N_0 - D_0L/\Delta} \right]^{-1}
\]
Depth Resolution Decreases with Distance

- The amount of depth information available falls off with the square of the object distance.
  - Solving for a measured displacement gives,
    \[
    \Delta = \frac{D_0 L}{(M_0 - M)f + D_0/N_0}.
    \]
  - As \( M \) decreases, \( \Delta \) rapidly approaches its limit of \( D_0 L/(M_0 f + D_0/N_0) \).
  - The rate of change in \( \Delta \) with \( A \),
    \[
    \frac{\partial \Delta}{\partial A} \approx -\frac{f^2 D L}{A^2 C^2} \quad \rightarrow \quad \frac{\partial \Delta}{\partial A} \approx -M^2 N^2 \frac{L}{D}.
    \]
Spatial Resolution and Pixel Size

- Spatial resolution is limited to the total number of pixels $mnk^2$.
- In order to achieve redundancy, the local magnification factor is set to $N < 1$.
  - Spatial resolution is reduced by $1/N^2$.
  - The total recoverable resolution is $\approx mnk^2N^2$
- Example: A $16 \times 16$ array of $0.5\mu m$ pixels with a magnification factor of $N_0 = 1/4$ produces a maximum resolution 16 times greater than the aperture count and 16 times lower than the pixel count.
• The minimum spot size for a diffraction limited system is approximately $\lambda/NA$. 
Spot Size Comparison

- The minimum spot size for a diffraction limited system is approximately $\frac{\lambda}{NA}$.

- The minimum useful pixel pitch is half the spot size using Rayleigh criterion.
Spot Size Comparison

- The minimum spot size for a diffraction limited system is approximately $\lambda/NA$.

- The minimum useful pixel pitch is half the spot size using Rayleigh criterion.

- Disparity from a Multi-Aperture system gives displacement which can be smaller than diffraction limit.
Multi-Aperture Color System

- Spectral separation by aperture
- No color contamination from neighboring pixels
- Facilitates the use of large dielectric stack height which allows high logic density
Projected Color Channels

- Color channels only overlap in the space above the detector
Conclusion

- CMOS technology is used to create multiple cameras on a chip.
- Depth map is extracted by solving the correspondence problem between multiple views of the same points in the primary focal plane.
- The amount of depth resolution available increases with decreasing pixel size while the 2D spatial resolution remains limited.
- The sensor architecture may be useful in improving the performance of color imaging by employing a per-aperture color filter.