Leveraging Pixel Value Certainty in Pixel-Shift and Other Multi-Shot Super-Resolution Processing

Henry (Hank) Dietz

COIMG-142: 9:10 – 9:30 AM, January 24, 2024

University of Kentucky
Electrical & Computer Engineering

Our Parsek Software Tool

- A **parsec** is approximately 3.26 light years: the distance at which a tiny angle offset of one arcsecond (1/3600º) spans one AU
- **Parsek** is Probabilistic Alignment Raw Stitcher Experiment from Kentucky, using tiny-offset captures to create a huge image
  - Primarily for stitching ***pixel shift images***
  - All merging of image data is based on ***pixel value confidence***
  - ***Only raw, uninterpolated, sensel values are used*** – even if the input is a JPEG, raw values are reconstructed
  - ***C++ code*** using **OpenCV** and **LibRaw** `unprocessed_raw`

Aggregate.Org UNBRIDLED COMPUTING

electronic IMAGING 2024 JANUARY 21–25

University of Kentucky

Multi-Shot Methods to Increase Resolution

- How much overlap between shots?
  - Small degree of overlap
  - Large degree of overlap
- What moves?
  - Camera and lens as a unit
  - Film/sensor only
- How does it move?
  - Manual positioning
  - Controlled motion system

Stitched Panoramas

- How much overlap between shots?
  - Small degree of overlap
  - Large degree of overlap
- What moves?
  - Camera and lens as a unit
  - Film/sensor only
- How does it move?
  - Manual positioning
  - Controlled motion system

Photo: NodalNinja.Com

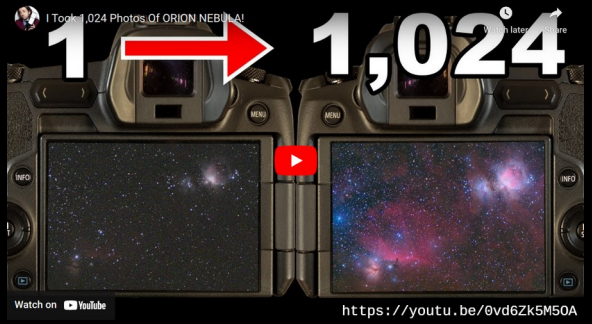
Photo: GigaPan.Com

Scanning Cameras and Scanning Backs

- How much overlap between shots?
  - Small degree of overlap
  - Large degree of overlap
- What moves?
  - Camera and lens as a unit
  - Film/sensor only
- How does it move?
  - Manual positioning
  - Controlled motion system

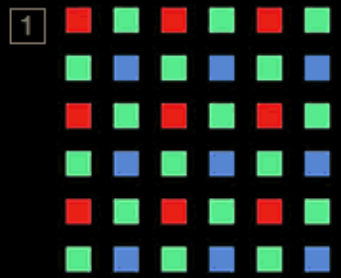
Image Stacking and Night Sight

- How much overlap between shots?
  - Small degree of overlap
  - Large degree of overlap
- What moves?
  - Camera and lens as a unit
  - Film/sensor only
- How does it move?
  - Manual positioning
  - Controlled motion system



Pixel Shift and Handheld Pixel Shift

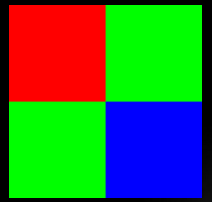
- How much overlap between shots?
  - Small degree of overlap
  - Large degree of overlap
- What moves?
  - Camera and lens as a unit
  - Film/sensor only
- How does it move?
  - Manual positioning
  - Controlled motion system

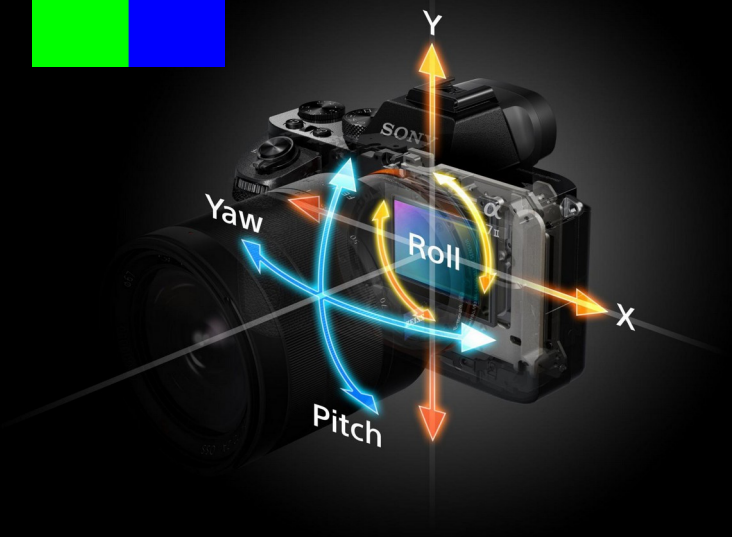


Animation: Olympus-IMS.com

Pixel Shift and Handheld Pixel Shift Implementation

- Uses IBIS: In-Body Image Stabilization
- Pixel shift uses just X, Y
  - 4 shots: (0,0), (1,0), (0,1), (1,1) improves color, noise
  - 16 shots: uses 1/2-pixel offsets improves color, noise, resolution
- Handheld Pixel Shift tries to keep offsets WRT original shot





The Primary Problems with Pixel Shift

- Scene content change during exposure sequence
- Accuracy of X, Y offsets
  - Camera + lens could move due to external or induced vibration (movement can include X, Y, Z, roll, pitch, and yaw)
  - Precision and accuracy of sensor motion control
- Sampling issues lower contrast
  - Anti-alias (AA) filter or poor lens reduces contrast at pixel level
  - What is the pixel fill factor (taking microlenses into account)?

Scene Content Change

- Temporal changes really can't be "fixed"
  - Detect by area average difference between images
  - Credible repair by either blurring region or excluding image data that differs from a selected reference
- Parsek probabilistic filtering methods:
  - Prefer values from low-detail regions (approximate motion blur)
  - Prefer values similar to reference (first listed) image
  - "Outlier" filtering prefers the most similar values
    - Set with > 50% of confidence sum
    - Largest set of similar values

Accuracy of X, Y Offsets

- Sony A7RV on a “solid tripod” for shots of Yosemite, here are the X, Y pixel offsets computed by alignment:

```
 0          0          -1.03392   -0.699499
-0.556742  -1.01952     -0.345406  -0.200595
-0.597450  -0.423268    -0.0492523 -1.371310
 1.11865   -1.11554      1.18047    0.114805
-0.160752  -0.393202    -0.202039  -1.49176
 0.859977  -1.86519      0.196493  -0.645636
-0.740545  -0.387017    -0.622772  -1.32642
 0.454802  -1.61298      0.576902  -0.915504
```

- These do **NOT** match the 0.5-pixel multiples intended
- Error from vibration, IBIS, imperfect alignment computation

Sampling Issues

- Sampling function is not precisely known
  - Anti-alias (AA) filter or poor lens reduces contrast at pixel level
  - What is the pixel fill factor (taking microlenses into account)?
- Resulting image needs to be sharpened to restore local contrast by Octave, Deblur, Unsharp mask, Multiscale, …
- Parsek demosaicing directly uses confidences, but also should be sensitive to patterns (like VNG, PPG, AHD, and AMaZE)

Pixel Shift, PixelShift2DNG vs. Parsek (A7RV)

```
 1.000000  0.000000   0.000000   1.000000  6.701e-5 -1.03392
 0.000000  1.000000   0.000000  -6.701e-5  1.000000 -0.699449

 1.000000  5.021e-5  -0.556742   1.000000  1.648e-5 -0.345406
-5.021e-5  1.000000  -1.01952   -1.648e-5  1.000000 -0.200595

 1.000000 -4.787e-5  -0.59745    1.000000 -6.402e-5 -0.0492523
 4.787e-5  1.000000  -0.423268   6.402e-5  1.000000 -1.37131

 1.000000 -4.063e-5   1.11865    1.000000 -3.182e-5  1.18047
 4.063e-5  1.000000  -1.11554    3.182e-5  1.000000  0.114805

 1.000000 -4.795e-5  -0.160752   1.000000 -5.590e-5 -0.202039
 4.795e-5  1.000000  -0.393202   5.590e-5  1.000000 -1.49176

 1.000000 -9.350e-5   0.859977   1.000000 -2.073e-5  0.196493
 9.350e-5  1.000000  -1.86519    2.073e-5  1.000000 -0.645636

 1.000000  3.040e-5  -0.740545   1.000000  2.817e-5 -0.622772
-3.040e-5  1.000000  -0.387017  -2.817e-5  1.000000 -1.32642

 1.000000 -9.104e-7   0.454802   1.000000 -4.782e-5  0.576902
 9.104e-7  1.000000  -1.61298    4.782e-5  1.000000 -0.915504
```

Note: Sony A7RV ISO100 16-shot pixel shift, Tamron 28-200mm @28mm;
largest movement should be ~0.006mm

Aggregate.Org UNBRIDLED COMPUTING

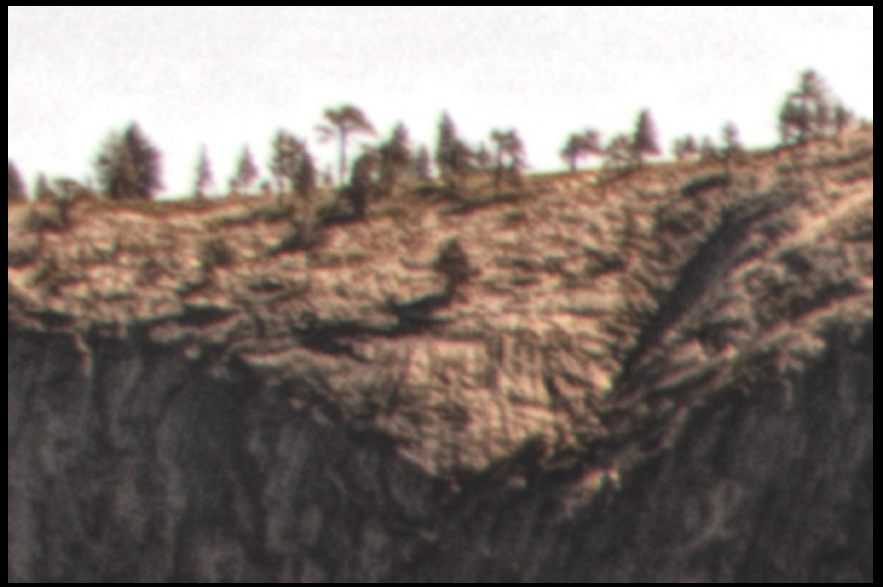
electronic IMAGING 2024 JANUARY 21–25

University of Kentucky

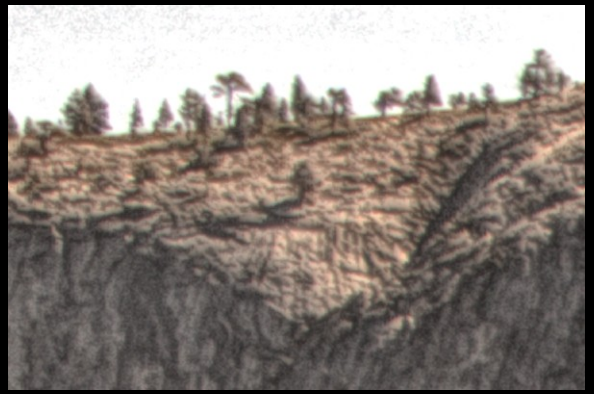
Pixel Shift Scaling Using Parsek (A7RV)

- Parsek can *render at **any** resolution*...

550MP **3X**:



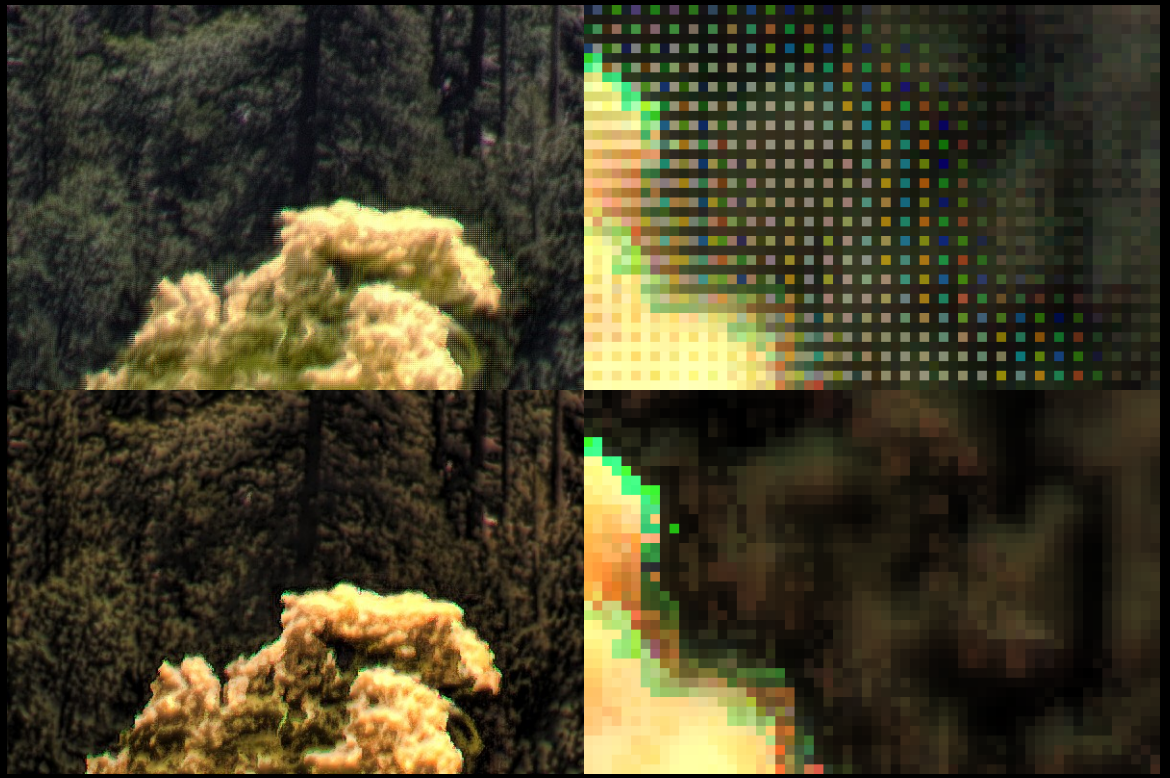
244MP **2X:**



61MP
1X:

Pixel Shift with Movement, PixelShift2DNG vs. Parsek

- Movement between shots makes different colors sample different scene content – a “dot grid” pattern results
- Parsek confidence is used to favor locally-consistent content



Application to Handheld with IBIS (A7II)

- Small camera + lens movements:
  - Pitch → Y, Yaw → X, Z is negligible
  - Roll needs matrix correction
- Example is 4 handheld shots, matrices:

```
 1.000000  0.000000   0.00000    1.000000 -0.000170  17.1762
 0.000000  1.000000   0.00000    0.000170  1.000000 – 1.77194

 0.999999  0.001518  12.0452     0.999998  0.001798 – 6.46442
-0.001518  0.999999  18.9449    -0.001798  0.999998  29.2788
```

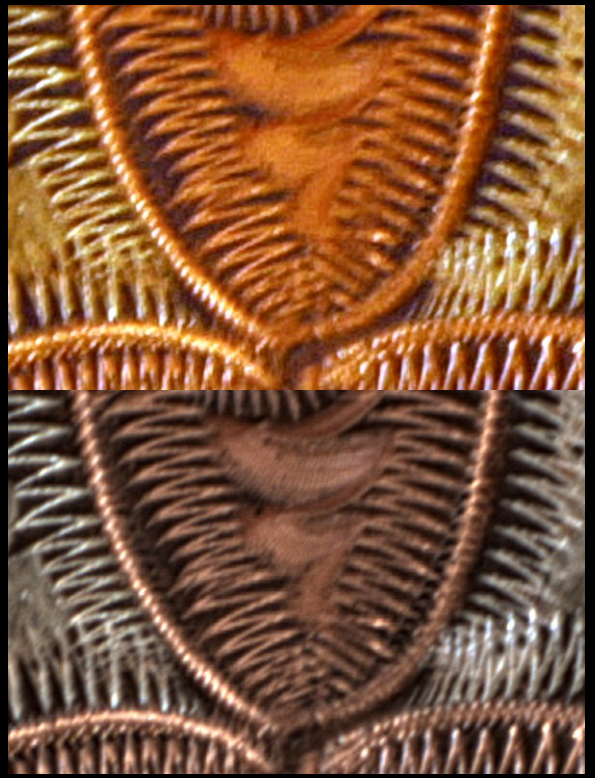
Note: Sony A7II ISO3200 **IBIS on** burst, Tamron 28-200mm @28mm;
worst-case movement < 0.2mm

Application to Tripod Mechanical Shutter (NEX5)

- Example is 7 tripod shots, matrices:

```
 1.000000  0.000000  0.00000    1.000000  0.000114 -0.199308
 0.000000  1.000000  0.00000   -0.000114  1.000000  1.63148

 1.000000 -0.000411  1.71246    1.000000 -0.000103  2.43466
 0.000411  1.000000 -3.31291    0.000103  1.000000 -3.31291

 1.000000  0.000163  1.33720    1.000000 -0.000289  2.79297
-0.000163  1.000000 -3.91932    0.000289 1.0000000 -5.00340

 1.000000 -0.000192  2.88670
 0.000192  1.000000 -6.75450
```

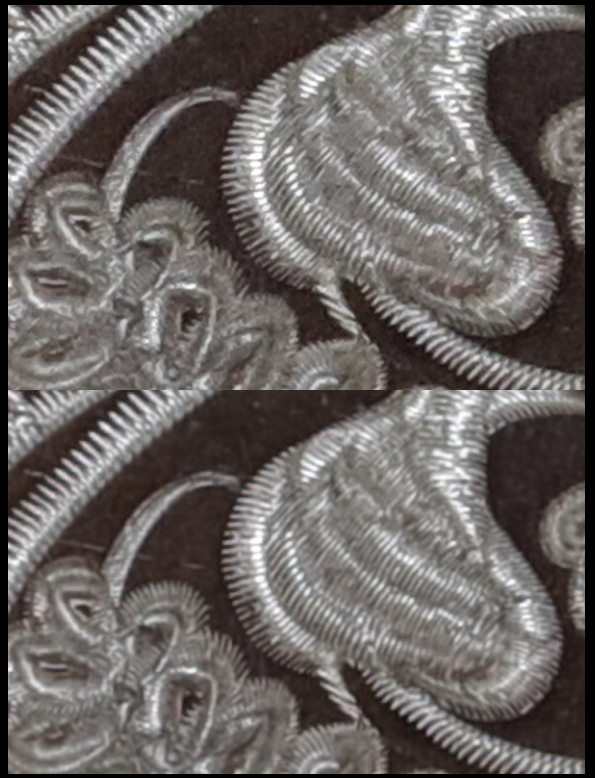
Note: Sony NEX5 ISO200, Takumar 35mm; movement < 0.04mm



Application to Tripod DSLR (5DIV)

```
 1.000000  0.000000  0.00000    1.000000  7.683e-5 -1.07518
 0.000000  1.000000  0.00000   -7.683e-5  1.000000  0.40627

 1.000000 -0.000168 -0.836035   1.000000 -0.000160 -1.05267
 0.000168  1.000000 -0.052061   0.000160  1.000000  0.140288

 1.000000 -0.000213  4.8045     1.000000 -0.000135 -2.21256
 0.000213  1.000000 -1.6784     0.000135  1.000000 -0.296225

 1.000000 -0.000178 -1.88743    1.000000 -3.481e-5 -3.07847
 0.000178  1.000000  0.288268   3.481e-5  1.000000 -0.849969

 1.000000 -1.785e-5 -1.93008    1.000000  1.228e-5 -2.30608
 1.785e-5  1.000000 -1.17242   -1.228e-5  1.000000 -1.25281

 1.000000  1.830e-5 -2.45083    1.000000 -0.000126 -2.73266
-1.830e-5  1.000000 -0.26872    0.000126  1.000000  0.27872

 1.000000 -6.387e-5 -2.42884    1.000000  1.228e-5 -1.38742
 6.387e-5  1.000000  1.29444   -1.228e-5  1.000000 -0.473685

 1.000000 -0.000269  2.13191    1.000000 -8.787e-5  0.539208
 0.000269  1.000000 -0.286397   8.787e-5  1.000000 -1.08307
```

Note: Canon 5DIV ISO1250, Yongnuo 50mm; movement < 0.03mm



Aggregate.Org UNBRIDLED COMPUTING

electronic IMAGING 2024 JANUARY 21–25

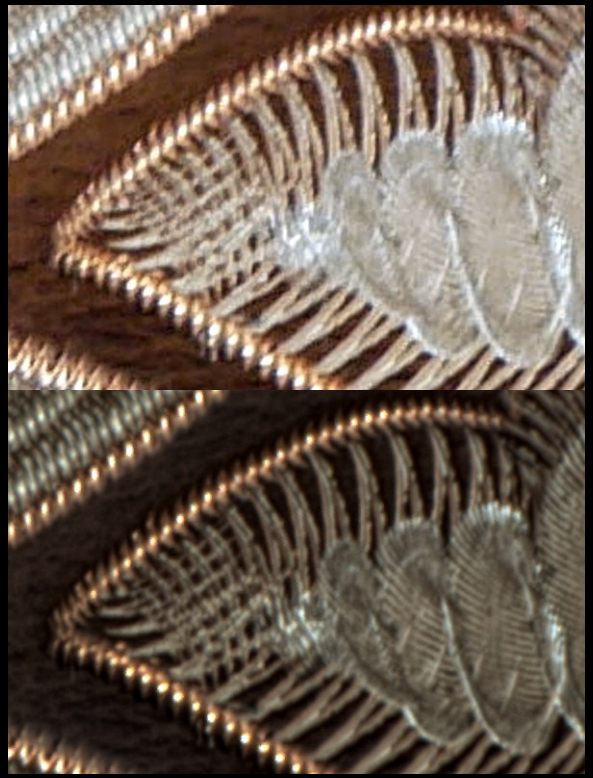
University of Kentucky

Application to Tripod with E-Shutter (GX850)

- Example is 10 tripod shots, matrices:

```
 1.000000  0.000000  0.00000    1.000000  4.251e-5 -1.09719
 0.000000  1.000000  0.00000   -4.251e-5  1.000000 -0.292483

 1.000000  9.998e-5  2.22964    1.000000  6.072e-5 -2.15389
-9.998e-5  1.000000 -0.624429  -6.072e-5  1.000000 -1.37466

 1.000000 -3.025e-5 -1.05566    1.000000  1.083e-5 -0.586833
 3.025e-5  1.000000 -1.44086    1.083e-5  1.000000  0.970092

 1.000000  4.965e-5 -0.918101   1.000000  3.201e-5 -1.36666
-4.965e-5  1.000000  0.803148  -3.201e-5  1.000000  0.112376

 1.000000  8.480e-5 -1.91083    1.000000  6.147e-5 -1.90292
-8.480e-5  1.000000  1.58552   -6.147e-5  1.000000  1.37994
```

Note: Lumix GX850 ISO200 **e-shutter**, Minolta 45mm; movement < 0.01mm



Parsek Command Line Options & Defaults

```
Usage: parsek -options input_files
-a int    set number of alignment iterations (50)
-c str    2x2 RGB CFA pattern is RGGB
-e flt    alignment termination epsilon (1e-12)
-f flt    filter pixels differing from first by more than (0.25)
-m flt    set maximum confidence (4)
-n        toggle neighborhood filtering (OFF)
-o str    set output file name (parsek.png)
-O flt    filter outliers differing by more than (0.25)
-p flt    prune image if alignment rotation/scaling exceeds (0.010000)
-r        toggle input images are raw (OFF)
-s        toggle smoothing in 24bpp to raw conversion (OFF)
-v        increment verbosity of messages
-X int    set final image int times input image x (columns); 2 by default
-x int    set final image x as int (columns)
-Y int    set final image int times input image y (rows); 2 by default
-y int    set final image y as int (rows)

Input images are converted to 16bpp raw; result is 16 bits per channel RGB
```

Conclusions – using Parsek

- Using **only the raw-sampled image data** improves stitch quality
- **Confidence-based merging reduces artifacts**
  - Weighting by precise spatial relationships
  - Allows simple filtering to reduce temporal artifacting
- **Real-world multi-shot capture has positional noise**
  - Tripod e-shutter pixel-shift does **NOT** move exactly as specified
  - Tripod typically constrains Y and Roll tighter than X
  - Handheld IBIS bursts gave +/- ~0.1mm, significant Roll
  - Tripod single-shot between +/- ~0.01mm and +/- ~0.05mm