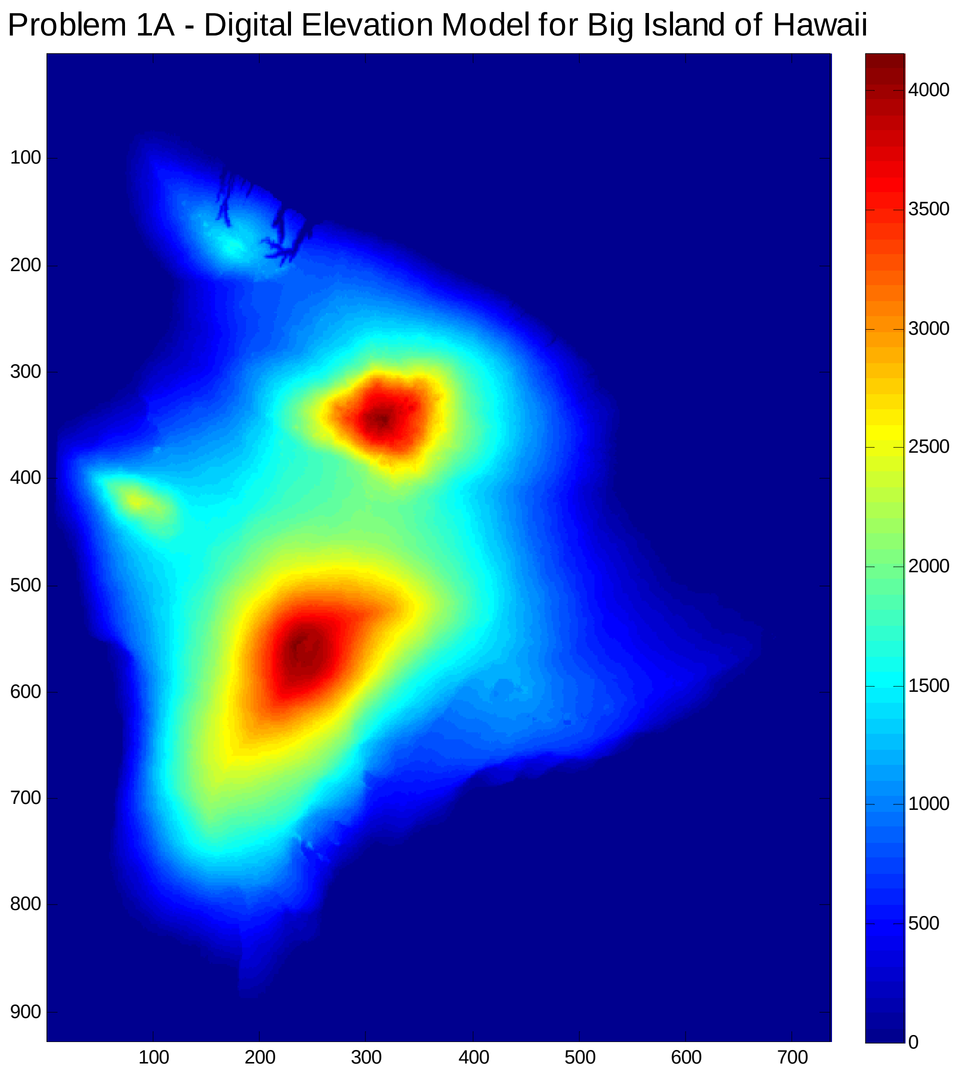
Christopher Tsai
December 26, 2007
EE 262 – 2D Imaging

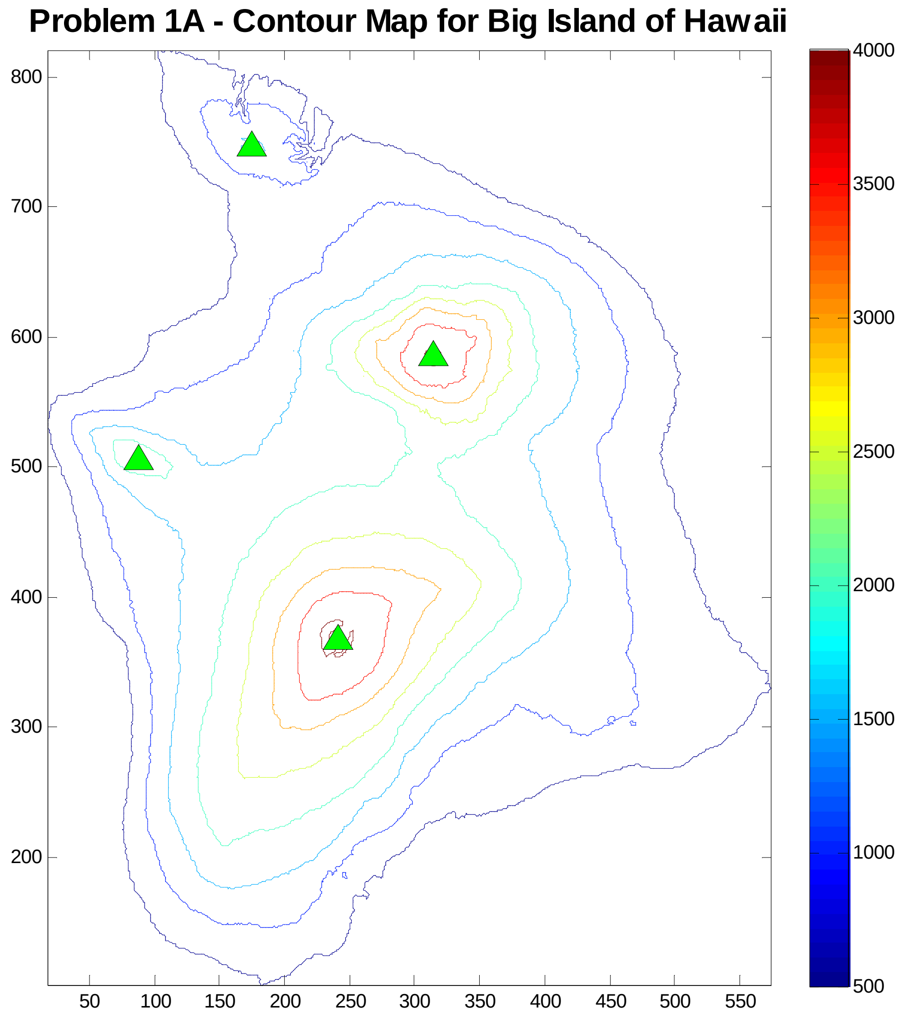
Problem Set II – Coherent and Incoherent Imaging

Problem #1 – Digital Elevation Model

We first generate a digital elevation model (DEM) of the big island of Hawaii to examine the topography cursorily:



The two larger red spots represent Hawaii's tallest peaks, Mauna Kea in the north and Mauna Loa in the south. To quantify the heights more precisely, however, we must plot the contours:



Designating each local maximum by a small green triangle, we count a total of *four* distinct peaks from the contour. If we know the pixel spacing in the image to be 180 m, then we can compute the distance between Hawaii's tallest peaks, Mauna Kea and Mauna Loa, from their DEM displacement:

$$\text{Physical Distance} = \text{Pixel Spacing} \cdot \sqrt{(x_{Loa} - x_{Kea})^2 + (y_{Loa} - y_{Kea})^2}$$

$$\text{Physical Distance} \approx 180 \text{ m } \cdot 227.112 \text{ pixels}$$

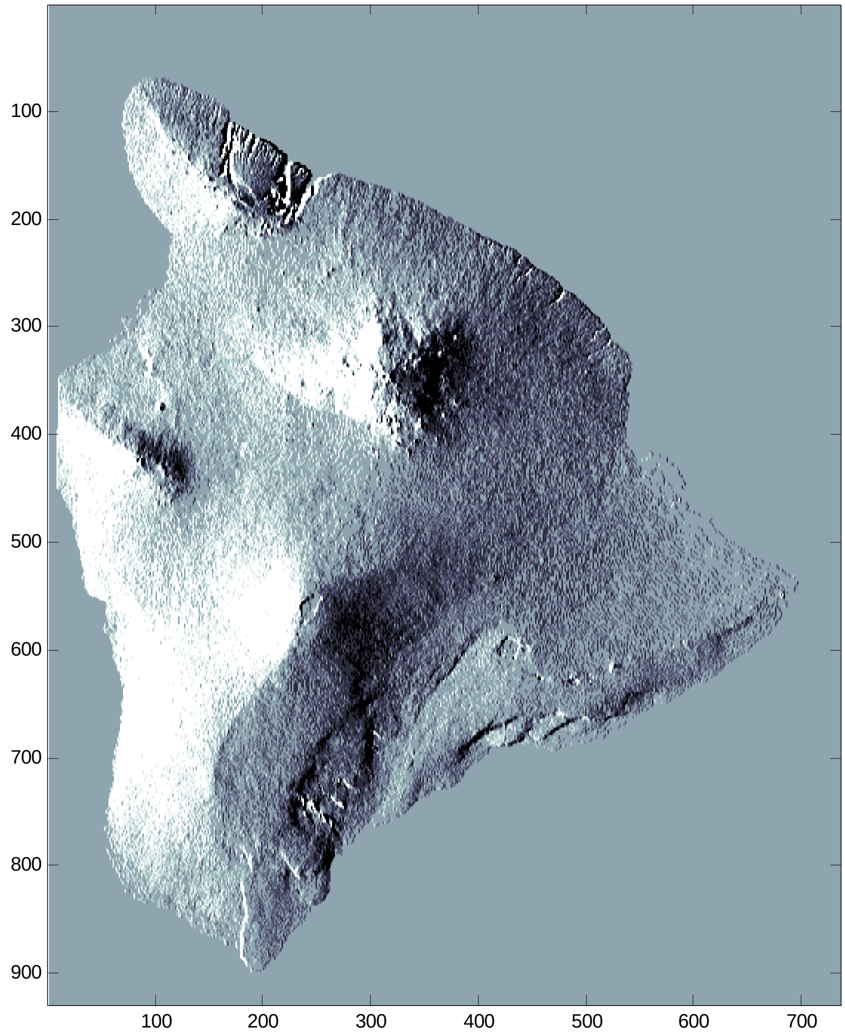
$$\boxed{\text{Physical Distance } \approx \mathbf{40,880.141\ m.}}$$

If we wish to model illumination of the island from the left, then we must locally simulate the difference equation for the original DEM image *f* over all neighboring rows *x* and columns *y*:

$$Left\ Illuminate\ \{f(x,y)\} = [f(x,y) - f(x-1,y)] + [f(x,y-1) - f(x-1,y-1)] + [f(x,y+1) - f(x-1,y+1)].$$

Solving this difference equation amounts to two-dimensional convolution with the positive-left negative-right shading kernel: $\begin{bmatrix} +1 & -1 \\ \boxed{+1} & -1 \\ +1 & -1 \end{bmatrix}$. The shaded relief image that results is

Problem 1B - Left-Illuminated Shaded Relief DEM of Big Island of Hawaii



Similarly, if we seek to model illumination from the image bottom, we must solve

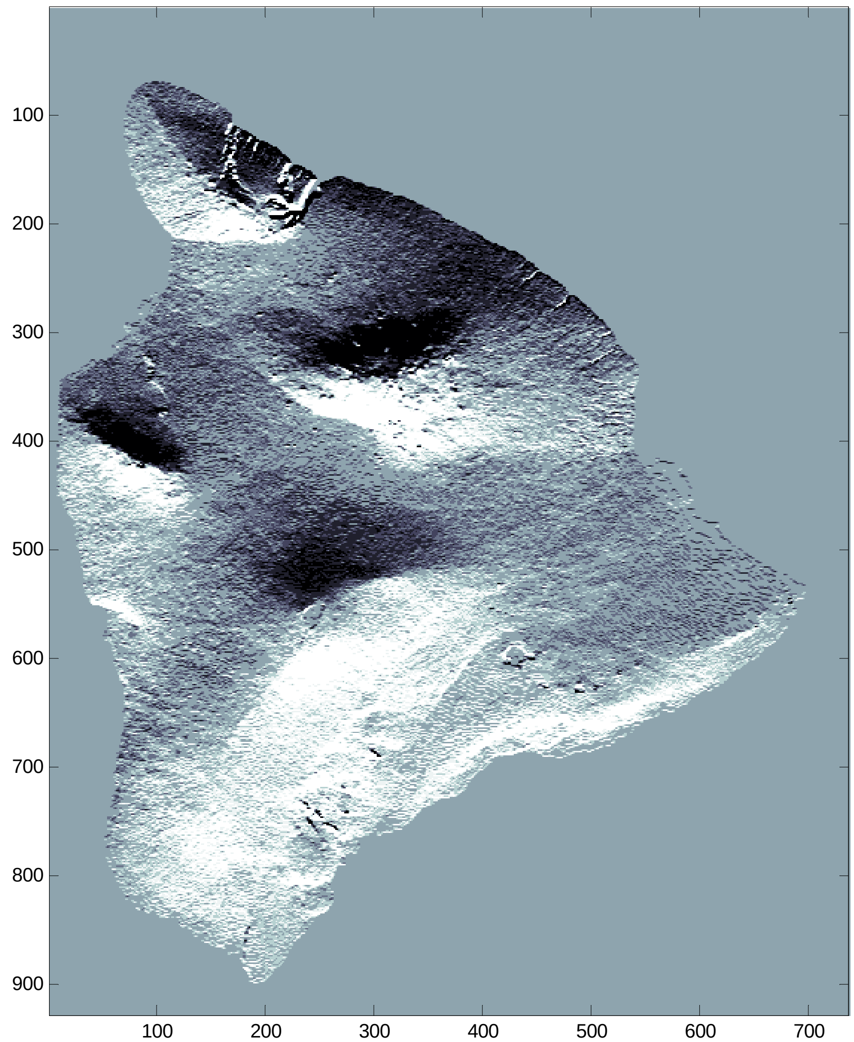
$$\textit{Lower Illuminate}\,\{f(x,y)\} = [f(x,y) - f(x,y-1)] + [f(x-1,y) - f(x-1,y-1)] + [f(x+1,y) - f(x+1,y-1)].$$

Solution amounts to two-dimensional convolution with the lower-positive upper-negative kernel:

$$\begin{bmatrix} -1 & -1 & -1 \\ +1 & \boxed{+1} & +1 \end{bmatrix}$$

…where the boxed element represents the center of convolution, and the ordered pair (x, y) represents (column index, row index). We produce the shaded relief image:

Problem 1C - Lower-Illuminated Shaded Relief DEM of Big Island of Hawaii



Problem #2 – Perspective Projection

In order to perform perspective projection, we effectively remove the height variable from consideration, normalizing by *(y + D)* in our initially trivariate coordinate system, where our chosen vantage point is (*D, H*) = (3000 m, 1200 m), representing the origin of our modified perspective. Essentially, we collapse the third dimension into the first two. We transform variables:

$$u = \frac{Dx}{y + D}$$

$$v = \frac{D(z - H)}{y + D} + H$$

We scan through the perspective plane variables according to our data matrix size:

$$x \in \left[-\left\lfloor\frac{N_{rows} - 1}{2}\right\rfloor, \left\lfloor\frac{N_{rows}}{2}\right\rfloor\right]$$

$$y \in \left[-\left\lfloor\frac{N_{columns} - 1}{2}\right\rfloor, \left\lfloor\frac{N_{columns}}{2}\right\rfloor\right]$$

According to the perspective transformation, the transformed variables span the ranges:

$$u \in \left[\frac{D \min x}{\min y + D}, \frac{D \max x}{\min y + D}\right] = \left[\frac{-D\left\lfloor\frac{N_{rows} - 1}{2}\right\rfloor}{-\left\lfloor\frac{N_{columns} - 1}{2}\right\rfloor + D}, \frac{D\left\lfloor\frac{N_{rows}}{2}\right\rfloor}{-\left\lfloor\frac{N_{columns} - 1}{2}\right\rfloor + D}\right]$$

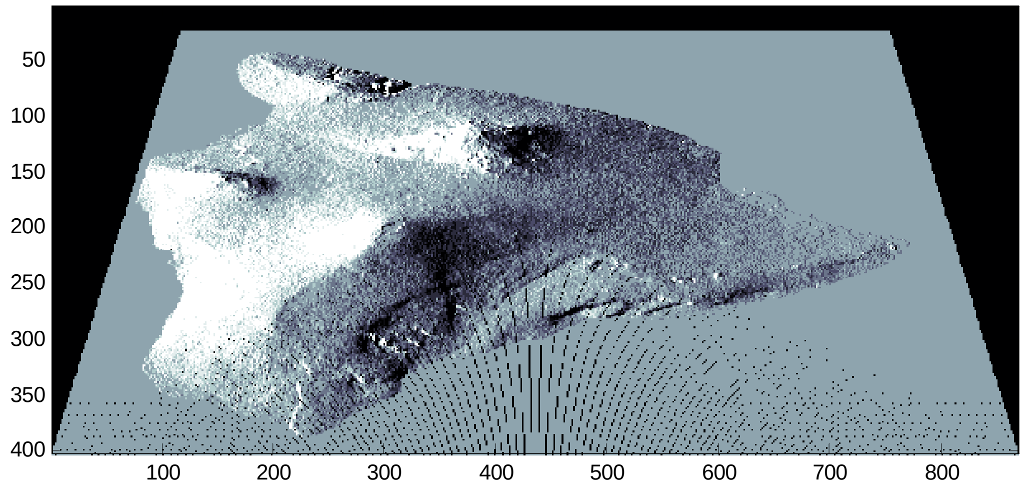
$$v \in \left[\frac{D(\min z - H)}{\min y + D} + H, \ \max\left\{\frac{D(\max z - H)}{\min y + D}, \frac{D(\max z - H)}{\max y + D}\right\} + H\right]$$

In discretizing our projection plane, we must increment both our original variables and transformed variables one unit at a time. We perform a shift of original image indices according to our newly rounded image indices:

$$\begin{cases} i_0 = \max y - y + 1 \rightarrow i = \max v - v + 1 \\ j_0 = x - \min x + 1 \rightarrow j = u - \min u + 1 \end{cases}$$

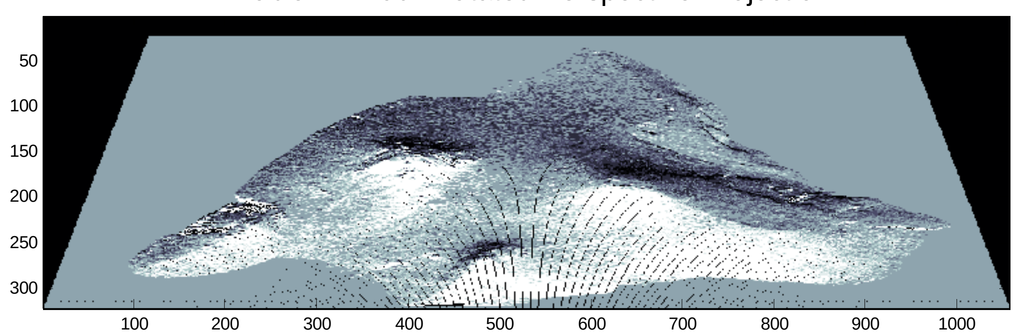
Upon converting our transformed coordinates (*u, v*) into array indices, we can plot the matrix on the perspective plane from two different vantage points:

Problem 2 - Unrotated Perspective Projection



We can rotate the perspective plane by 90° simply by pre-multiplying the coordinates by a rotation matrix. This linear transformation results in the rotated image:

Problem 2 - 90°-Rotated Perspective Projection



Notice that both images exhibit the wounds of transformation: black holes or stripes of seemingly disparate data. These empty points are simply the vestige of an imperfect injective mapping, as some points map more densely into the perspective plane than others; consequently, some areas receive fewer mappings by virtue of their magnification from our particular perspective, which obviously favors the proximity to the horizon, where the need for image detail decreases. All in all, though the array is rectangular, the illusion of perspective in our image actually stretches the proximity relative to the distant points, so that the lower half of array receives greater emphasis in the image, leading to a seeming point deficiency.

Problem #3 – Pinhole Camera Equations

Using the pinhole camera equations with a vertical displacement to the pinhole, we perform the following modified change of variables:

$$u = -\frac{D_2 x}{y + D_1}$$

$$v = \frac{D_2(z - H)}{y + D_1} + H$$

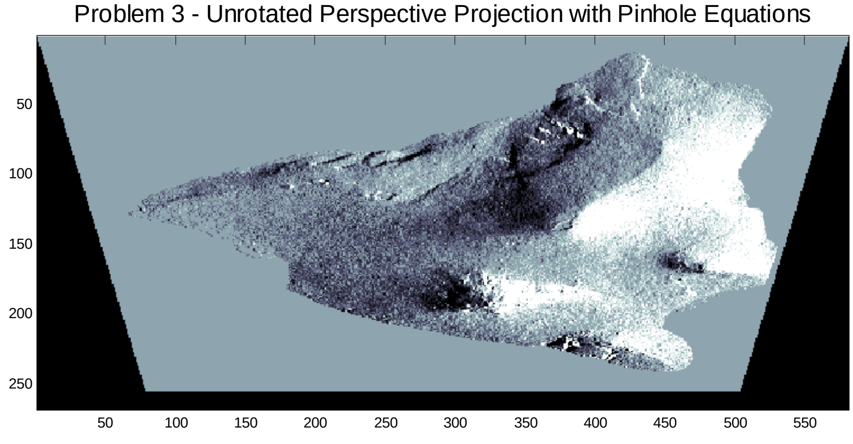
where D_2 represents observer displacement from the *y*-axis, and D_1 represents the pinhole height, independent of the observer position. Thus, we position the pinhole at $(x, y, z) = (0, 0, D_1)$ and the observer at $(x, y, z) = (0, -D_2\,, H)$.

$$u \in \left[-\frac{D_2 \min x}{\min y + D_1}, \qquad -\frac{D_2 \max x}{\min y + D_1}\right]$$

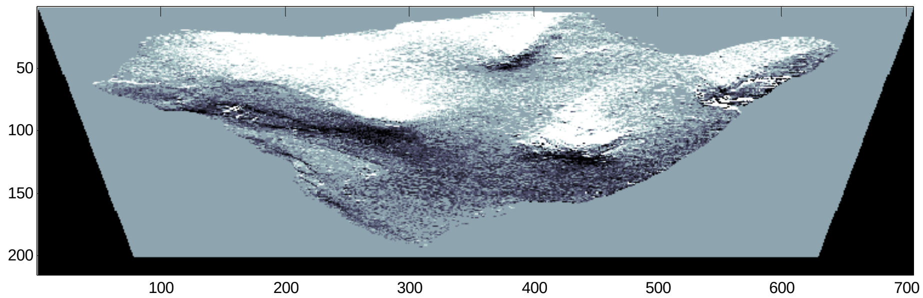
$$v \in \left[H - \max\left\{\frac{D_2(\max z - H)}{\min y + D_1}, \frac{D_2(\max z - H)}{\max y + D_1}\right\}, \;\; H - \frac{D_2(\min z - H)}{\min y + D_1}\right]$$

We generate our shaded relief images using the same iterative discretization of transform variables, rotating with the same rotation matrix. However, this set of transform equations yields a different-looking image. Unsurprisingly, the pinhole camera images appear inverted in the image plane, because the pinhole image appears inverted following passage:

Problem 3 - Unrotated Perspective Projection with Pinhole Equations



Problem 3 - 90°-Rotated Perspective Projection with Pinhole Equations

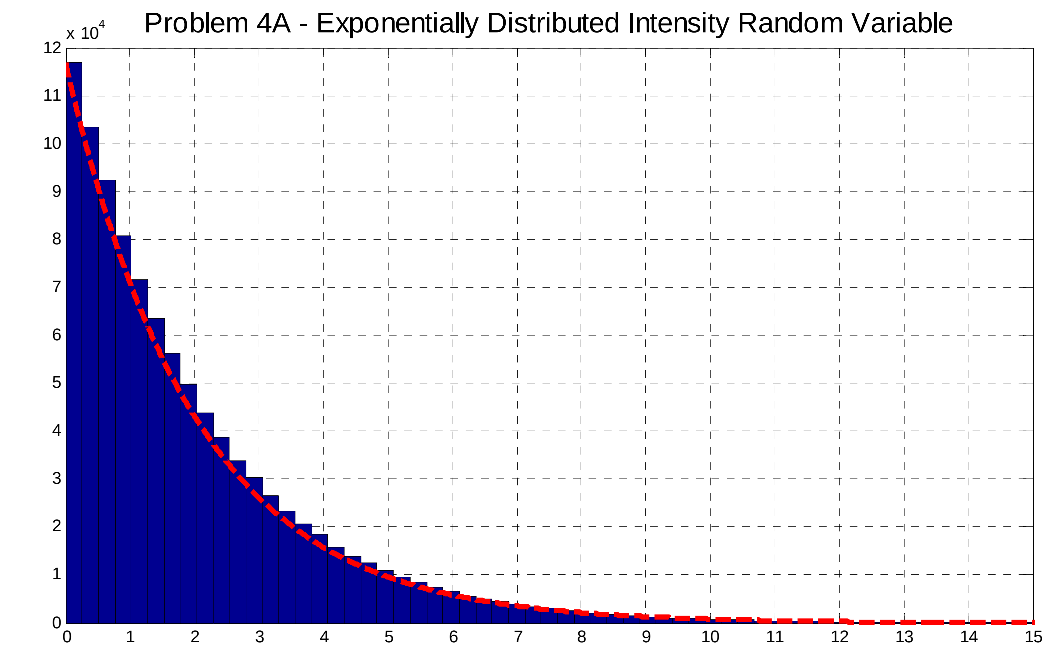


Problem #4 – Gaussian Random Number Generator

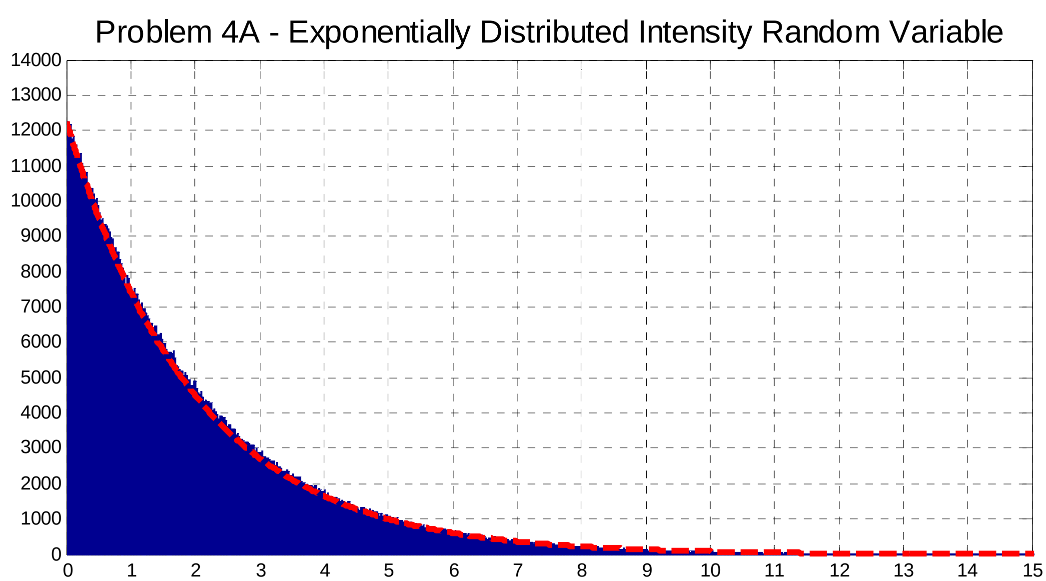
A uniform random variable ~$U\left[-\frac{1}{2}, +\frac{1}{2}\right]$ has zero mean and variance $\sigma^2 = \frac{1}{12}$. Thus, we can approximate a Gaussian random vector by summing twelve vectors of these uniform random variables. Since expectation is linear, the Gaussian mean will remain zero, while the variance of the sum of independent random variables is the sum of the variances, thus elevating our Gaussian variance to unity. We populate an entire vector with these independent sums, and hence produce a string of normally distributed random numbers with zero mean and unit variance. We combine independently drawn pairs in (C = X + *j*Y) to produce complex Gaussian random variables. To ensure that our complex combination's constituents are successfully Gaussian, we measure their means, variances, and third moments across vectors of length 1,000,000 (one million iterations), as tabulated below:

Random Variable	**Mean μ**	**Variance σ^2**	**Third Moment**
Real Part {C} = X	-0.000380 ≈ **0**	0.999081 ≈ **1**	0.000197 ≈ **0**
Imaginary Part {C} = Y	-0.000349 ≈ **0**	0.997051 ≈ **1**	-0.000846 ≈ **0**

We proceed to combine complex Gaussian random variables incoherently by generating a sequence of random intensity values I = $|C|^2 = C\bar{C}$, which possesses an exponential distribution. We juxtapose our probability density function (histogram of I) with the theoretical exponential distribution $f_I(r) = \begin{cases} \lambda e^{-\lambda r} & for\ r \geq 0 \\ 0 & for\ r < 0 \end{cases}$ with parameter $\lambda = \frac{1}{2}$:



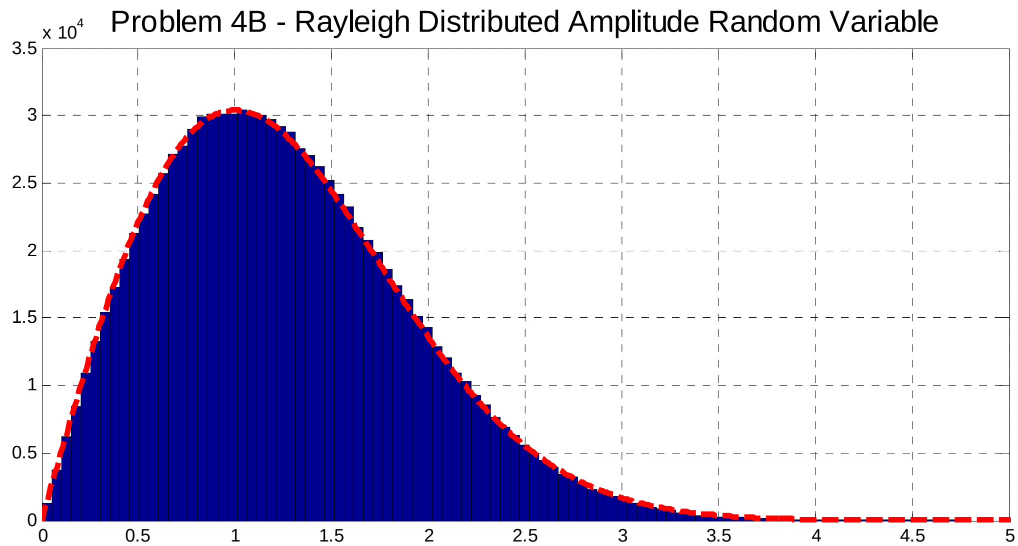
By decreasing the bin width, we can approximate the ideal exponential distribution (dotted line) even more closely, as we aver below:

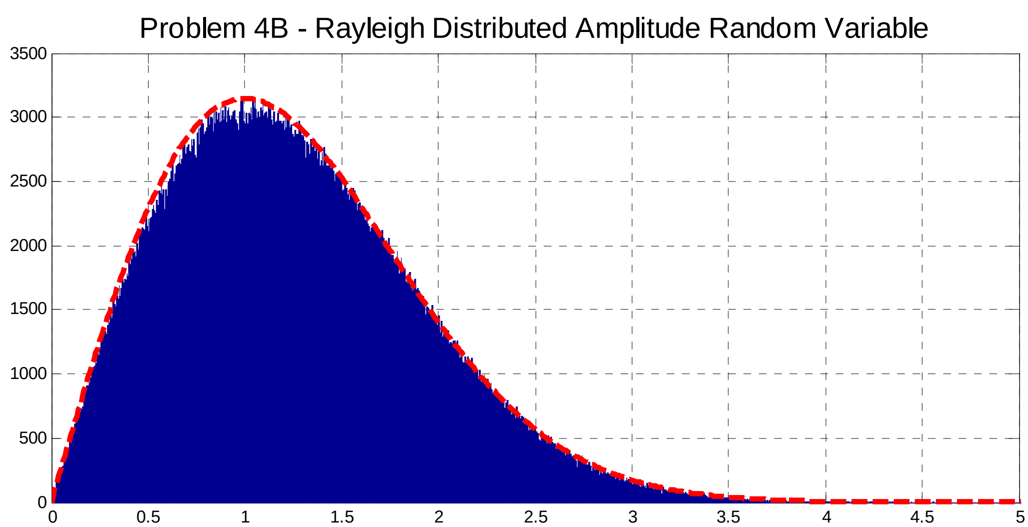


	Mean μ	**Variance σ^2**
Experimental Distribution of I	$1.996130 \approx \mathbf{2}$	$3.781063 \approx \mathbf{4}$
True Exponential Distribution	$\frac{1}{\lambda} = \frac{1}{1/2} = 2$	$\frac{1}{\lambda^2} = \frac{1}{1/4} = 4$

Thus, the theoretical distribution matches the experimentally-inferred values closely! With even more random variables, we can approximate the exponential distribution arbitrarily accurately.

If we generate a sequence of amplitudes $A = |\mathrm{C}| = \sqrt{\mathrm{C}\bar{\mathrm{C}}}$, then the resultant distribution is the classic Rayleigh distribution instead, as we ascertain by juxtaposing the empirical and theoretical distributions $f_{\mathrm{A}}(a) = \begin{cases} \frac{a}{\sigma^2} e^{-\frac{a^2}{2\sigma^2}} & for\ a \geq 0 \\ 0 & for\ a < 0 \end{cases}$ with parameter $\sigma^2 = \frac{4-\pi}{2}$ once again:

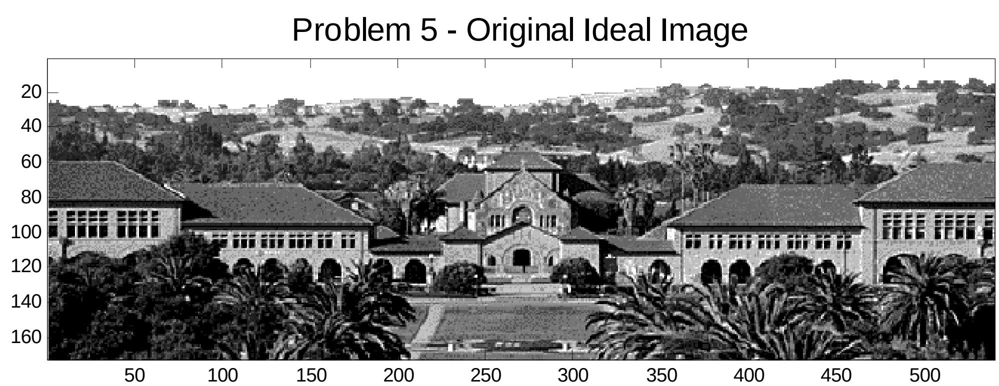


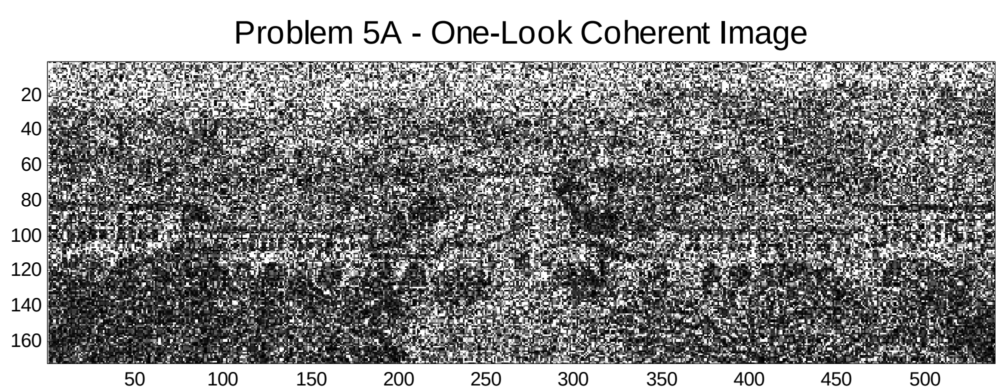


	Mean μ	**Variance σ^2**
Experimental Distribution of A	1.256170	0.418167
True Rayleigh Distribution	$\frac{1}{2}\sqrt{\frac{\pi}{\lambda}} = \frac{1}{2}\sqrt{\frac{\pi}{1/2}} = \sqrt{\frac{\pi}{2}} \approx 1.253314$	$\frac{4-\pi}{4\lambda} = \frac{4-\pi}{4(1/2)} = \frac{4-\pi}{2} \approx 0.429204$

Problem #5 – Coherent vs. Incoherent Imaging

We simulate a one-look coherent image by generating a matrix of complex Gaussian random variables C = X + *j*Y; in this complex image, the real and imaginary parts X and Y are normally distributed with zero mean and variance $\frac{I}{2}$, where I represents the pixel value intensity as prescribed in the input image1. In other words, we modulate the complex Gaussian random vector with the scalar constant $\sqrt{\frac{f}{2}}$, where the image value *f(x,y)* changes from pixel to pixel to reflect the original ideal structure in the picture of the Stanford quad. We plot the image intensity, $|C|^2$ and balk at the dominance of the coherent speckle; barely any detail is perceptible:

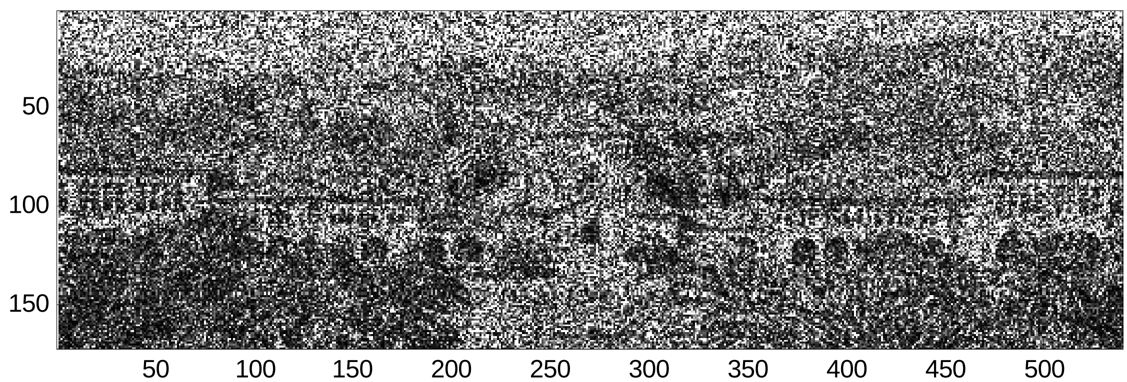




However, we can improve the resolution of the image and penetrate the speckle if we average several random draws in a multi-look image. Performing ten looks, we view the improved image:

Problem 5 - Original Ideal Image

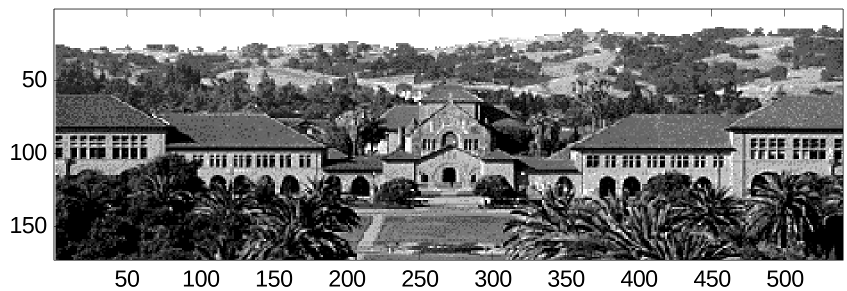
Problem 5A - One-Look Coherent Image



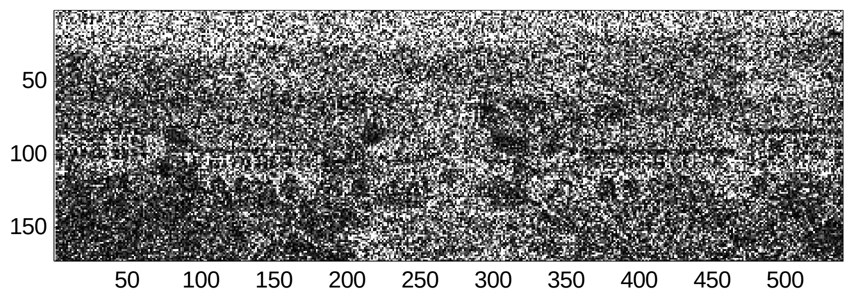
Problem 5B - Ten-Look Coherent Image

With ten looks, the detail in the Stanford image begins to materialize, with the windows, rooftops, arches, and trees entering the realm of recognizance. However, some grainy speckle still remains to a small degree, corrupting the visibility of the distant hills and enchanting the image with an artificial air. Nevertheless, the ten-looked image much more closely resembles the original image than the one-look rendition; the features now enjoy much more definite shape, boasting higher signal-to-noise visibility. We persist in our strategy, taking one hundred looks and redisplaying the image:

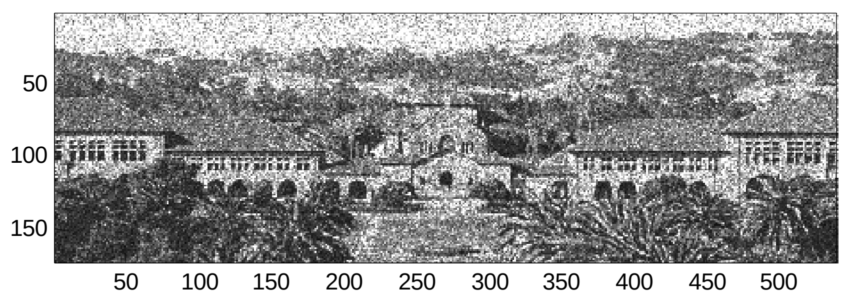
Problem 5 - Original Ideal Image



Problem 5A - One-Look Coherent Image

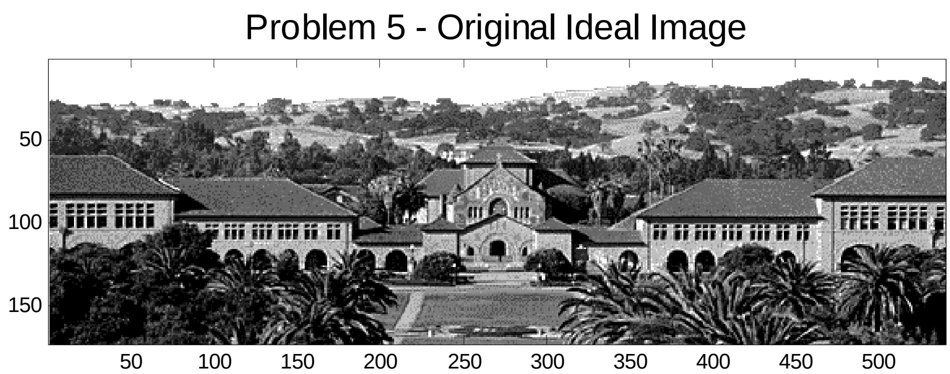


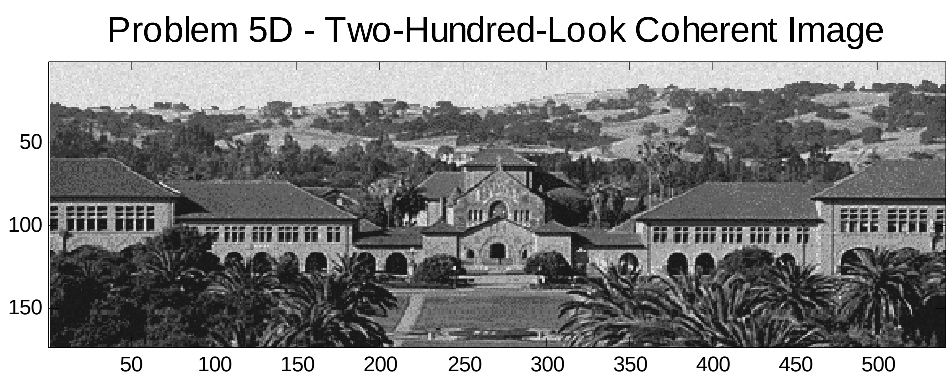
Problem 5B - Ten-Look Coherent Image



Problem 5C - Hundred-Look Coherent Image

The image now boasts better contrast, but some gray speckle adulterates the skies, so we try again. With two hundred draws, we see slight improvement in the spotted skies:

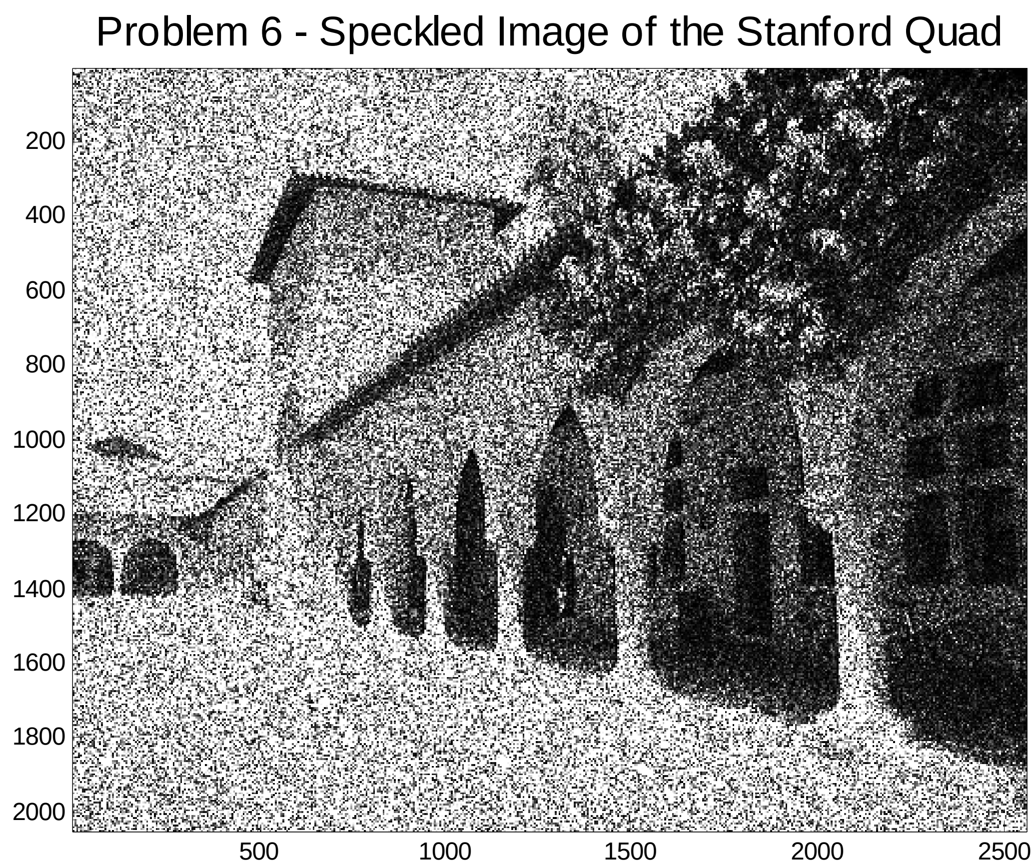




However, despite the general quality throughout the image, the discontinuous variations in sky color render the coherent image distinct. In order to match the original ideal image, we must remove all noise, which is possible only with an infinite number of looks. In other words, because a finite average never completely removes the randomness of speckle noise, only an infinite number of looks can perfectly match the original image. However, without prior knowledge of coherent imaging, the flawed human eye, bound by the limits of vision, cannot distinguish extremely subtle differences, so we can achieve a *working* match with 200 looks, with an image juxtaposed below. As theory intimates, we can still discern speckle dots in the transparent (nearly white) skyline, where small gray points are marginally visible to the trained eye. However, these grains do not inordinately debase the quality of the image, so we consider this a working average, more of a satisfactory heuristic and subjective choice than any mathematical axiom.

Problem #6 – Despeckling an Image

A coherent camera begot the following image, resulting in high speckle content and low visibility:



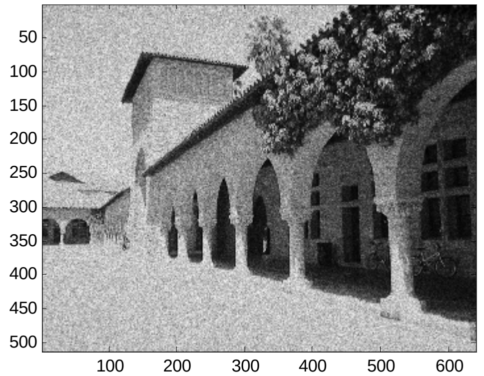
We can reduce the speckle in this image by averaging pixels, obtaining two looks in each direction. As a byproduct, we decrease the original image size fourfold, but the resolution improves:

Problem 6A - Despeckled Two-Look Image of the Stanford Quad

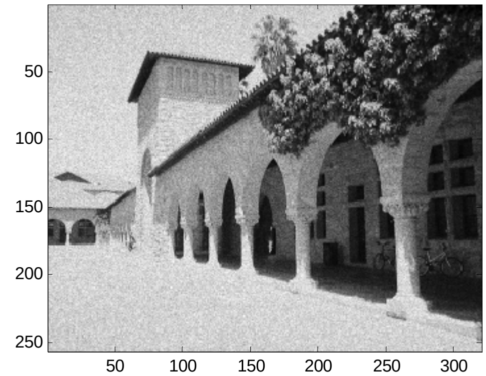
The image still suffers from grain, but the reduction in the number of specks opens the door for improved perceptibility, especially concerning details in the image edges, such as the arches of the arcade. All in all, the smooth regions of the image still bear noticeable grain, but we can better discern image details such as the palm tree, column friezes, and building edges.

Proceeding to minimize the image even further through arithmetic averaging, we assay a variety of look sizes:

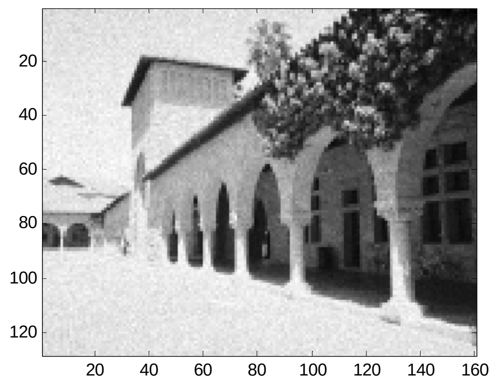
Problem 6C - Despeckled Four-Look Image of the Stanford Quad



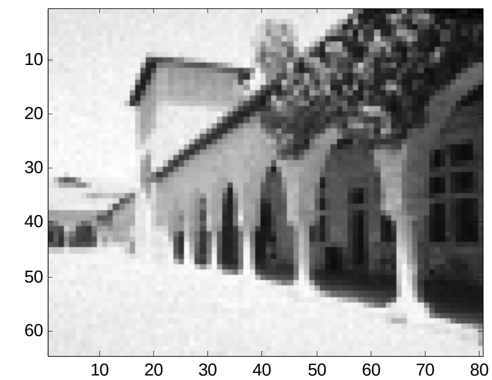
Problem 6C - Despeckled Eight-Look Image of the Stanford Quad



Problem 6C - Despeckled Sixteen-Look Image of the Stanford Quad



Problem 6C - Despeckled Thirty-Two-Look Image of the Stanford Quad



Of all of our trials, the eight-look image most discreetly balances detail and smoothness. The four-look image still shows too many grains to merit consideration as a "despeckled" image. Meanwhile, the six-teen look and thirty-two look images exhibit the flaws of superfluous averaging; the edges of image structures cluster together in blocks, the artifact of blurring and smearing of too many pixels to the point that insufficient data remain to reconstruct objects faithfully. Thus, despite marginally noticeable speckle in the eight-look image, the eightfold averaging has not reduced detail beyond the point of human visual tolerability.