# CS 6475 Final Exam

**Due** Dec 9 at 5pm **Points** 100 **Questions** 34
**Available** Dec 2 at 5pm - Dec 9 at 5pm 7 days **Time Limit** 150 Minutes
**Allowed Attempts** 2

## Instructions

**Computational Photography**

**Final Exam**

Please read the following instructions carefully.

- You will have 2.5 hours to answer the questions.
- Except the first and the last question (which are worth 0.5 each), T/F questions are worth 1 point, and all other question are worth between 2 and 4 points
- Questions can be of multiple-choice, true/false, matching or fill-in-the-blank type. All your answers will be graded automatically. You will get some feedback at the end of the exam.

**Important Note:** If for some reason you get kicked out of your exam, try logging back into Canvas -->Quizzes and try to resume the exam. If you run into any other technical difficulties during the exam, send the Instructors a message on Piazza and we'll try to resolve your issue.

Best of luck!

Take the Quiz Again

## Attempt History

| | Attempt | Time | Score |
|---|---|---|---|
| **LATEST** | Attempt 1 | 144 minutes | 69.78 out of 100 |

Correct answers are hidden.

Score for this attempt: **69.78** out of 100
Submitted Dec 7 at 5:46pm
This attempt took 144 minutes.

### Question 1 0.5 / 0.5 pts

I certify that:

- [x] I am talking this exam solely and entirely on my own, without any help from any other individual.
- [x] I am the student who is enrolled in this class
- [x] I am aware of the Georgia Tech Honor Code (link (http://www.policylibrary.gatech.edu/student-affairs/academic-honor-code) ) and I affirm it here, as I take this exam.
- [x] I will NOT print or save any part of this exam, for any purpose whatsoever.

Thanks.

## Question 2

2 / 2 pts

[CP02a-2] Consider an CMYK Image where each channel is 8 bits. The resolution of this image is height = 3456 and width = 5184. Assuming that 1 Kilobyte = 1024 bytes, what is the exact memory requirement of this image?

- ○ 8192 Megabytes
- ◉ 69984 Kilobytes
- ○ 52488 Kilobytes
- ○ 4299816960 bits

YES!

W x H x BitsPerPixelPerChannel x Number of Channels / 8192

8192 is the number of bits in a kilobyte.
8 Bits/Pixel
4 Number of Channels for CMYK

Partial

## Question 3

3.2 / 4 pts

[CP02b-2] Which of the following is an accurate description of an Image Histogram?

- [ ] It plots the number of intensities for each pixel value.
- [x] Photographers can use them as an aid to decide whether image detail has been lost to over or under exposure.
- [ ] Should not ever be applied to subregions of images separately.
- [ ] Evaluating exposure saturation from histograms is less useful when using a raw image format (which captures the actual sensor responses) as the dynamic range of the displayed image may only be an approximation to that in the raw file.
- [x] Histograms show the intensity distribution of an image.
- [x] Can be separate for each channel.
- [x] It plots the number of pixels at each intensity value.

Incorrect

## Question 4

0 / 3 pts

[CP02c-2] Below is the function describing the "Overlay" blend mode:

$$f_{blend}(a,b) = \begin{cases} 2ab, & \text{if } a < 0.5 \\ 1-2(1-a)(1-b), & \text{otherwise} \end{cases}$$

Mark all of the true statements about the "Overlay" blend mode from the list below.

- [x] It combines the "Multiply" and the "Screen" blend modes depending on the pixel value of the bottom layer.
- [x] It is the reason we see the green effect in the lecture videos.

  Nope that is due to "Darken" Please review Lecture "Digital Images" or Lecture 02-3

- [ ] The parts of the top layer where the base layer is light become darker, and the parts where the base layer is dark become lighter.
- [ ] It models the Dodge blend mode, well-known by dark room photographers.

  Review Lecture "Digital Images" or Lecture 02-3

## Question 5

3 / 3 pts

[CP02d1-1] Given the two attached images from the lectures, please explain the differences between them in terms of how they were generated from 3

different images.

Hint: Review the lectures where these images were used!

Image DS-1:

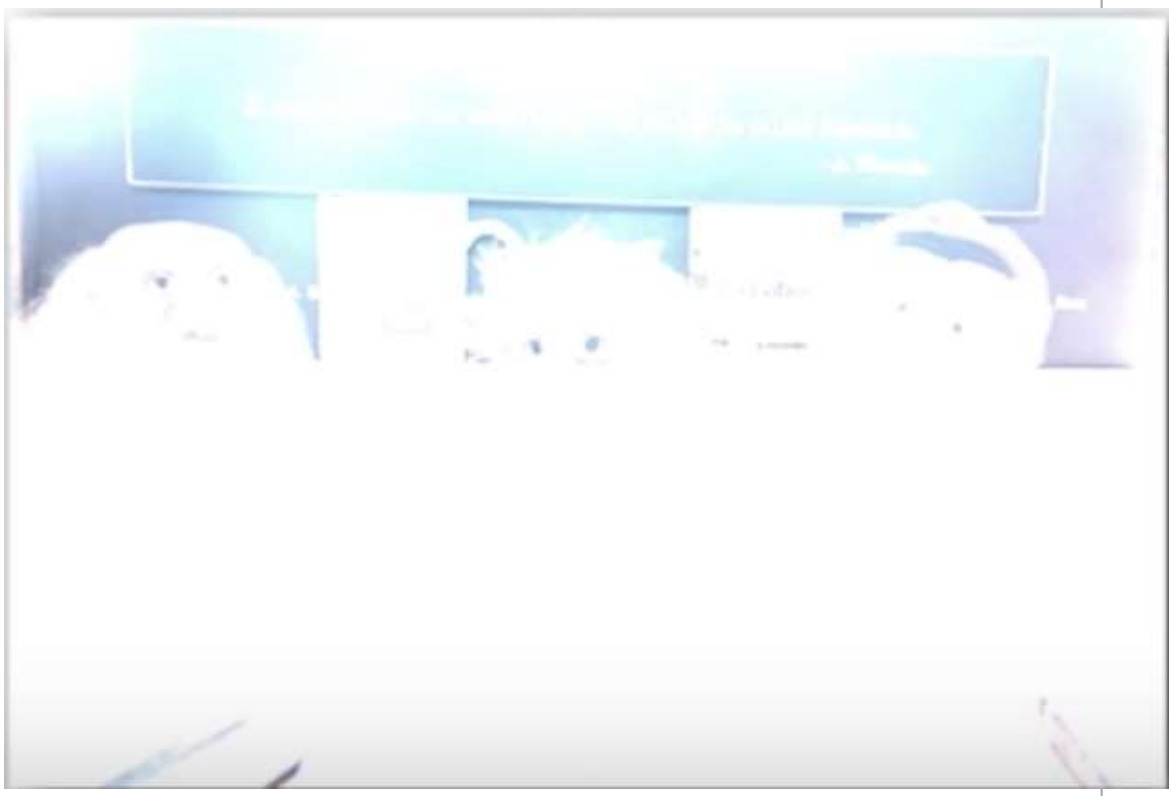

Image DS-2:

DS-1 image an image without any Bayer corrections, while DS-2 is after Bayer correction has been applied.

DS-1 image is 3 images added without any correction, while DS-2 is the same 3 images added with 1/3 weight factor (ie., 0.33*I1 + 0.33*I2 + 0.33*I3).

DS-1 image is an overexposed image, while DS-2 is NOT.

DS-1 image is 3 images added without any correction, while DS-2 is DS-1 with histogram equalization.

## Question 6

1 / 1 pts

[CP02d2-3] Converting images from integer to floating point data types completely eliminates the possibility of arithmetic overflow and underflow in computational photography.

True

False

Correct! All fixed-width precision data types suffer from over and underflow, but converting to float is usually good enough

## Question 7

3 / 3 pts

[CP02e1-1] Select all the correct statements below.

Convolution is...

- [x] equivalent to cross-correlation when the kernel is both horizontally and vertically symmetrical.
- [x] Commutative: $F * G = G * F$
- [ ] a measure of similarity of two waveforms.
- [x] an operation that calculates the area of overlap between two functions.

## Question 8

3 / 3 pts

[CP02e2-1] Select all the choices below which are correct given the following equation.

$$G[i,j] = \frac{1}{(2k+1)^2} \sum_{u=-k}^{k} \sum_{v=-k}^{k} F[i+u, j+v]$$

- [ ] This is the equation for convolution with uniform weights over a neighborhood of pixels, where *k* is the size of the neighborhood (above, below, left, right).
- [ ] This equation only applies Gaussian kernels, as weights are distributed across the neighborhood.
- [x] This is the equation for cross-correlation with uniform weights over a neighborhood of pixels, where k is the size of the neighborhood (above, below, left, right).
- [ ] This is the general form of an equation for convolution over a neighborhood of pixels, where *k* is the size of the neighborhood (above, below, left, right).

- [x] This equation only applies to square or average smoothing, as weights are equally distributed across the neighborhood.

Incorrect

## Question 9

0 / 3 pts

[CP02g-1] Select all of the true statements about image gradients from the list below.

- [x] The Prewitt, Sobel, and Roberts kernels are used to reduce noise sensitivity in edge detection.

Reducing noise sensitivity in gradient calculations is not a significant feature of these operators.

- [x] Edges in images are usually aligned with the local gradient direction.

Edges are usually perpendicular to the gradient direction.

- [x] Gradient calculations are very sensitive to noise in the image.

- [ ] The method of finite differences is the only exact operator to calculate image gradients.

See lecture "Groups of Pixels" and Lecture 02-6 on Udacity

Partial

## Question 10

2.67 / 4 pts

[CP03a-2] Select all of the true statements about pinhole photography from the list below.

- [x] Usually suffers from low light because of the size of the opening / aperture.
- [x] Has infinite depth of field. Everything is in focus.
- [x] Ideally, has virtually no geometric distortion; straight lines remain straight.
- [x] Usually suffers from motion blur.

Partial

## Question 11

2.67 / 4 pts

[CP03b-2] Select all of the true statements about aperture from the list below. Recall that the aperture number N relates the area of an aperture opening A and the focal length f according to the equation:

$$A = \pi\left(\frac{f}{2N}\right)^2$$

- [ ] Doubling N reduces Area by 4 times, and therefore halves the light received.
- [x] The radius of the opening is simply f/2N (from the above equation of Area of the opening, which is a circle).
- [x] The amount of light that falls on a sensor or film in a camera is proportional to the area of the aperture opening, and is referred to as RADIANCE.

IRRADIANCE, not radiance. See lecture 03-3

- [x] Low f-number (N) on a lens usually means it has a BIG lens. This is especially noticeable for Telephoto lenses, which have larger focal lengths. [A 800mm lens of f-number 4, will have 200mm aperture diameter]

- [x] The aperture number, or the f-number (N) usually marked on all lenses, is designed to give irradiance irrespective of the lens in use.

## Question 12

4 / 4 pts

[CP03c-2] Select all of the true statements about lenses from the following list.

- [ ] Field of view (FOV) of a lens only depends on the Focal Length because the sensor is placed on the focal plane of the lens.

- [x] The Combined Focal length of a compound lens can vary and depends on the distance between the lenses and their individual focal lengths.

- [x] A lens only forms a focused image on a screen placed at the focal length of the lens, otherwise the image formed is unfocused and suffers from optical blur.

- [x] Focal length is a fixed parameter for simple lenses.

## Question 13

4 / 4 pts

[CP04a1-2] Choose all of the correct factors to consider to choose the optimal window size for image blending from the following list:

- [x] To avoid seams: Window = size of largest prominent “feature”
- [x] To avoid ghosting: Window ≤ 2× size of smallest prominent “feature”
- [ ] Image frequency content should occupy two pyramid levels
- [x] Largest frequency ≤ 2 × size of smallest frequency

Lecture 04-2.

## Question 14

3 / 3 pts

[CP04a2-1] Choose all of the statements that are CORRECT about a Laplacian Pyramid from the list below:

A Laplacian is simply computed using

- [ ] $L_k = \text{REDUCE}(g_{k-1})$

- [ ] Each Laplacian Image in the Pyramid is a multiplicative product of two consecutive levels of a Gaussian Pyramid

Each Laplacian is computed using

- [x] $L_l = g_l - \text{EXPAND}(g_{l+1})$
- [x] A Laplacian Pyramid is a series of “error” images, $L_0$, $L_1$, $L_2$, ...

Incorrect

## Question 15

0 / 4 pts

[CP04a3-1] Which of the following statements are TRUE for using Cuts vs. Blending images?

- [x] Using Cuts is better when there are too many objects in the image and registration is hard.
- [x] Using Cuts is better when there is motion that causes ghosting, as the same object will have moved to another point in the image during blending.
- [x] Seam Carving is not similar in terms of computation to Cuts. Cuts are used to merge two images. Seams remove a set of pixels from the same image, and the methods are completely different.

No! Seam Carving uses similar approach on the same image to find a "cut" to remove or add.

- [x] Cuts are like median filtering as they give you an exact pixel value from one of the images, whereas blending merges pixel values between images.

See Lecture 04-4

## Question 16

1 / 1 pts

[CP04a4-1] Higher layers in a Gaussian pyramid (i.e., smaller frames) contain the higher frequency components of the image in the frequency domain.

true

false

Correct! Successive applications of Gaussian (low-pass) filters progressively eliminates high frequencies by one octave per application, so the highest level of the pyramid contains the lowest frequency content.

## Question 17

3 / 3 pts

[CP04a5-1] In Laplacian pyramid blending, pixel intensity G(i, j) near the image borders is affected by the choice of boundary condition. Consider pixels padded to the left of the input image border. Burt & Adelson used _____, which is equivalent to assuming that the second derivative is zero.

G(i, j) = k (an arbitrary constant value)

G(-i, j) = G(i-1, j)

G(-i, j) = G(i, j)

None of the Above

G(-i, j) = 2G(0, j) - G(i, j)

## Question 18

4 / 4 pts

[CP04b1-1] Please select from the following characteristics of Good Features.

- [ ] Variability - Variety of metrics that define a feature.
- [ ] Dominant -- Give a strong response to x-correlation
- [x] Saliency/Matchability -- Distinctive description
- [x] Locality -- Relatively small area of the image; robust to clutter and occlusion
- [x] Repeatability/Precision -- Find the same feature despite geometric and photometric transformations

Partial

## Question 19

0.75 / 3 pts

[CP04b2-1] Which of the following is CORRECT about the Harris Detector?

- [ ] Harris detectors are NOT Invariant to Image Scale changes. One needs to use Fourier transforms to convert the image into the Frequency domain and model the variations.
- [ ] Harris detectors are NOT Invariant to Image Scale changes. One needs to use a scale-space representation using Pyramids to detect features at different scales.
- [x] Harris detectors are Invariant to Rotation.
- [ ] Harris detectors are Invariant to Translation.
- [x] Harris detectors are Invariant to Image Scale changes
- [x] Harris detectors are Invariant to Image Intensity Variations.

Partial

## Question 20

2.67 / 4 pts

[CP05a1-2]Which of the following statements are true about Projective Transformation?

- [x] 4 point correspondences needed for computation

  See Lecture "Image Transformations and Warping" or Lecture 05-1 on Udacity

- [ ] Only the top two columns of the transformation matrix need to be computed to model this transformation.

- [ ] All 8 values except the bottom right element (which should be 1) need to be computed

- [ ] 8 point correspondences needed for computation

- [ ] 6 Degrees of Freedom

- [x] 8 Degrees of Freedom

Partial

## Question 21

2 / 4 pts

[CP05b2-1] Determine which of the following statements are CORRECT about a stereo method to compute depth or disparity in a scene

- [x] The Epipolar constraint for computing disparity makes searching for corresponding pixels in a stereo pair easier.

- [x] A simple stereo system used to compute 3D scene geometry assumes that there are two cameras, just displaced slightly. No other information is needed about the cameras except the distance between them.

How about focal length? Need it for computing 3D, but not for disparity. See Lecture "Stereo Vision" or Lecture 05-5 on Udacity. Simple system also assume same focal length for both cameras.

- [x] The disparity computed from a stereo pair is usually larger for closer surfaces than farther ones. This is due to parallax in the scene.

- [ ] The Epipolar constraint does not provide any computational efficiency in the case of a Kinect RGBD sensor.

## Question 22

1 / 1 pts

[CP05b3-1] Planar projection can NOT be used to make panoramas from images taken by rotating a camera around a single axis (i.e., the camera must undergo horizontal or vertical translation).

- ( ) True
- (•) False

## Question 23

3 / 3 pts

[CP05b4-2] Because planar projection panoramas preserve straight lines, you only need a _____ transform to align the images.

- (●) Perspective
- ( ) None of the Above
- ( ) Affine
- ( ) Rotation
- ( ) Translation

Correct! Affine preserves straight lines, but also preserve parallel lines (which are not preserved in planar projection panoramas).

## Question 24

3 / 3 pts

[CP05b5-3] The HDR process described by Debevec & Malik requires at least _____ measurements to ensure a sufficiently overdetermined system of equations to recover the irradiance values and response curve. (Let N = number of pixels; P = number of images; $Z_{max}$ = max pixel value; $Z_{min}$ = min pixel value)

- ( ) None of the Above
- ( ) 255
- ( ) N * (Zmax - Zmin)
- (●) N(P-1) > (Zmax - Zmin)
- ( ) 256

Incorrect

## Question 25

0 / 1 pts

[CP05b6-2] The response curve relating the values recorded from a digital camera sensor to irradiance E is a linear function.

- True
- False

Review “Recovering High Dynamic Range Radiance Maps from Photographs” by Debevec & Malik. Pay attention to the difference between the response curve and physical sensor response.

## Question 26

2 / 2 pts

[CP06a-2] What is the approximate minimum video frame rate for humans to perceive flicker-free motion?

- 24 frames/second
- 25 frames/second
- 30 frames/second
- 29.97 frames/second
- 12 frames/second

## Question 27

4 / 4 pts

[CP06b-2] Select the statements from the following which are correct for the concept of Video Textures.

Crossfading and blending create too much blur and should never be used for video textures.

Video Textures can be applied to sub-regions of images, like video sprites.

The primary concept that supports Video Texture analysis is that similar frames repeat in videos.

Though L1 and L2 similarity metrics are discussed for generating video textures, other similarity metrics can also be used to compare frames.

Video Textures only work when there is an image texture in a video.

Partial

## Question 28

1.33 / 4 pts

[CP06c1-1] Which of the following statements is true about the Video Stabilization system discussed in lecture 06-3?

Cropping is used to crop the view, which avoids problems with a rolling shutter.

It is a 3D camera path stabilization method, where a 3D path is computed and then a smoothing process applied.

It is a 2D camera path stabilization method, where only estimates of 2D motion are used, then constrained using standard notions of camera movements like pan or dolly.

- [ ] Cropping is used to crop the view, which avoids dealing with hole filling. When the whole frame is aligned, we may see regions with no pixels from the original video as the camera shakes.

- [x] Rolling shutter adds unwanted non-rigid motion in the video due to a delay in readout from photosites.

- [ ] It is a 2D camera path stabilization method, where only estimates of 2D motion are used, then constrained using cropping.

- [ ] Rolling shutter can be removed by adding median filtering in time.

Partial

## Question 29

3 / 4 pts

[CP07b1-1] Which of the following statements are correct about Epsilon or Coded Photography?

- [ ] Coded Photography cannot be used to take a 'standard' picture

- [x] A coded aperture essentially changes the aperture to provide variations in a captured image, which can then be computationally adapted.

- [x] Coded Photography is akin to Bayer Patterns. It encodes a code with an image, which can be used to extract a novel image.

- [ ] Low light performance and image resolution are not affected by adding coding to apertures

Epsilon Photography assumes that multiple images are taken and then combined to form a novel image.

Coded photography uses a "code" to encode variations in an image (or video) in neighboring pixels. Epsilon photography captures all variations and tends to deal with it in a post-processing.

## Question 30

4 / 4 pts

[CP07b2-2] Which of the following statements are CORRECT about a Light field Camera?

A hologram is represented as a 7-D Light Field.

One can build a light field camera, capable of depth, focus, and illumination estimation by combining a lens and micro-lens array system.

Panoramas can't be represented using the plenoptic function.

A plenoptic or a light field camera attempts to capture a light field rather than pixels. The captured light field can be rendered as a traditional image in pixels as a post-capture step.

Incorrect

## Question 31

0 / 3 pts

[CP08-6] Consider the paper on "Poisson Image Editing" in module 08. Which of the following relates to that effort?

Spots and blemishes are removed from fur images by separating out the brightness component from details in a selected region and replacing the brightness by harmonic interpolation (solving a Laplace equation) of the brightness at the selection boundary.

Nope, this isn't related to Poisson editing.

The mathematical tool at the heart of the approach is the Poisson partial differential equation with Dirichlet boundary conditions which specifies the Laplacian of an unknown function over the domain of interest, along with the unknown function values over the boundary of the domain.

Using this approach allows transferring objects to a new scene without any need for precise delineation of object boundaries.

A system is introduced to edit an image via a sparse set of its edge elements (edgels).

Nope, this isn't related to Poisson editing.

See referenced paper from module 8

Partial

## Question 32

1 / 3 pts

[CP08-2] Select the statements that are correct for the "Interactive Photomontage" approach presented in 08

Alignment of images if NOT required for the processing of images.

Gradient-domain image fusion in the color space is used to align the colors amongst the stack of images.

Cuts are used to merge and generate a new image

Images are blended to generate a new image

It works on a stack of images, along the lines of Epsilon Photography

Partial

## Question 33

1.5 / 3 pts

[CP08-7] "PatchMatch" is a method widely used for image editing. Which of the following statements is true for this method? (see Module 08)

PatchMatch is an interactive structural image-editing framework, which makes it possible to grab portions of an image and move them around

PatchMatch is an algorithm for image completion, which lets a user simply erase an unwanted portion of an image, which the computer automatically synthesizes a fill region that plausibly matches the remainder of the image

The core of the PatchMatch system is an algorithm for computing patch correspondences efficiently using an approximate Nearest Neighbor approach.

## Question 34

0.5 / 0.5 pts

[Closing] Reminder and recertification on closing:

I certify that:

- [x] I took this exam solely and entirely on my own, without any help from any other individual.
- [x] I will not copy or print this exam for any reason!
- [x] I am the student who is enrolled in this class.
- [x] I am aware of the Georgia Tech Honor Code (link (http://www.policylibrary.gatech.edu/student-affairs/academic-honor-code) ) and I affirm it here, as I take this exam.

Thanks.

Quiz Score: **69.78** out of 100