Introduction to Computer Vision

Final Exam (Example)

- Answer the following 4 questions.
- You may write your answers either in Hebrew or in English.
- Length of exam: 3 hours.

Good luck!

Q1:
Let \( f(x, y) \) be an \( M \times N \) image, and let \( F(u, v) \) be its Fourier transform. Let \( g(x, y) \) be an image of dimensions \( (3M) \times (3N) \), whose Fourier transform \( G(u, v) \) is generated from \( F(u, v) \) by inserting two rows of zeros (0’s) between every two rows of \( F \), and two columns of zeros (0’s) between every two columns of \( F \):

\[
G(u, v) = \begin{cases} 
F(k, l) & u = 3k, \ v = 3l \\
& k = 0, ..., M - 1 \\
& l = 0, ..., N - 1 \\
0 & \text{otherwise}
\end{cases}
\]

What does the image \( g(x, y) \) look like in terms of \( f(x, y) \)? Show mathematically and explain the result in short. Remember that \( g(x, y) \) is of size \( (3M) \times (3N) \).

Q2:
1. How many point correspondences are needed to recover a homography between two images? Explain why.
2. 12 points are painted on a dark plane. 5 points are colored pure red, 5 are colored pure green, and 2 are colored white. Two cameras take pictures of the plane. One camera is equipped with a red filter (doesn’t see pure green color) and one camera is equipped with a green filter (doesn’t see pure red color). How many points does each camera see? Is this number sufficient in order to compute the homography relating the two images? Explain why.
Q3:
A camera is imaging an object at two time instances. Point correspondences are given across the two images. Can depth be recovered if:

1. There is only a camera rotation between the two images?
2. There is only an object rotation between the two images?

Q4:
We mentioned in class two problems that can be solved by SVD factorization: (1) Recovering of 3D structure and motion from a collection of 2D images, and (2) Photometric stereo (recovering shape and lighting from a collection of images).

Choose ONE of the above problems and answer the following questions:

1. What is the input matrix?
2. What are the underlying assumptions about the scene or the camera that allow you to use factorization?
3. What is the rank of the input matrix if these assumptions hold (assuming there is no noise)?
4. What kind of ambiguity remains after the SVD step? Explain in short how this ambiguity can be removed.