Progress Report

Bardia Mojra

April 18, 2022

Robotic Vision Lab

The University of Texas at Arlington

1 Specific Research Goals

- VPQEKF (RAL - April 1st): Work on the paper.
- DLO Manipulation Dataset (ICRA - September)

2 To Do

- QEKF Paper - 30% extension (April 1st):
  - Edit VEst section and add updates.
- QEKF/QuEst+VEst Implementation (Feb. 28th):
  - Implement QuEst 5-point: Done, debugging.
  - Feature point extraction: implement semantic segmentation
  - Implement VEst
  - Address scale factor (depth-scale) issues: DL solutions?
  - Address "hand off" issue when objects enter or leave field of view
  - Real-time streaming images for real-time operation (optional)
  - Experiments
  - Noise issue: noise cannot be modeled
- DLO Manipulation:
  - Related work literature review
  - Real dataset + paper (September 2022 - ICRA):
    - Design, discuss and build a data collection and test rig.
  - Unity dataset
    - Recreate virtual duplicates of physical test material
    - Model dynamics and deformity

3 Progress

The following items are listed in the order of priority:

- VPQEKF (RAL - April 1st, 2022): This week, I investigated discrepancies in QuEst Python implementation. I began with loading extracted and matched SIFT features from the Matlab implementation and comparing all computations line by line. I was able to narrow down the error to Scipy's *singular value decomposition* (SVD) routine. This particular SVD routine offers two LAPACK driver options, *'gesvd'* and *'gesdd'*. The 'gesdd' method is based on divide-and-conquer parallelization and is faster. The 'gesvd' method is the general rectangular approach and is also used by Matlab. I tested both options and got similar results. In both cases, the SVD produced a V matrix with values that matched those produced by Matlab up to 18 significant figures with an exception of its *Null Space*. The null space values matched only up to 1 or 2 significant figures and it alludes to accumulated rounding error and propagated observation noise. Before we discuss available remedies for decomposition null spaces, it is worth diving deep into vector spaces, linear algebra, geometric algebra, and singular value decomposition. But rather than going over the basics, I will highlight how they are connected by definition. In short, a *vector space* is a notation that represents some space on basis of *normal vectors*, e.g. i,j,k vectors in 3D Cartesian coordinate frame. These normal vectors correspond to x, y, and z axes which are orthogonal to each other by definition. These constraints form the *Orthonormal* constraint which is assumed by definition when dealing with linear and geometric algebras unless mentioned otherwise. The orthogonality can be confirmed by carrying out the inner product which returns zero confirming independence among axes. In a non-deterministic sense, it can be viewed as zero correlation among axis, but we will leave discussing the probabilistic side of this regime for another time (look into EKF for a more detailed example of such methods). At last, the SVD is used to decompose an often noisy observation into an optimally similar linear model. It is important to note that the SVD transformation is to the frequency domain, this may not be apparent to untrained eyes. By definition, the SVD is a generalized *Fast Fourier Transform* (FFT) and is widely used in computer vision as a tool for principal component analysis and data compression. To better understand the SVD operation of interest, I plotted its singular values and their cumulative sum as they show matrix rank and the distribution of information in the original matrix. The SVD returns U, S, and V which are the left complex unitary matrix, singular values, and right complex unitary matrix, respectively. U and V are complex skew-symmetric matrices and when

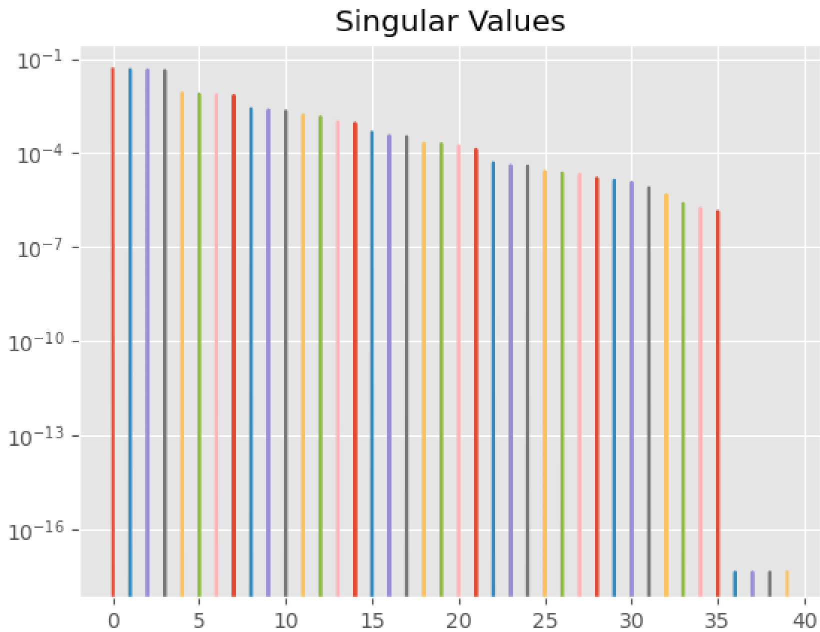

Figure 1: Singular Values

multiplied form an orthonormal coordinate frame with singular values representing scaler magnitude for vectors on the corresponding axis.

Back to discussing null space created in a situation where we aim to compute the coordinate transformation between two image frames. Given a set of matched correspondences, we aim to compute the closest linear transform between the two frames. Instead of using RANSAC which is too costly, we can simply compute the nearest orthogonal matrix. We know that the transformation is linear, the solution must be unique, and the axis must be orthogonal. Thus, this becomes an optimization problem where we seek to find the linear transform solution that yields the smallest Euclidean distance as it maps the initial coordinate frame to the latter. The Kabsch algorithm (also known as Wahba's problem) claims to provide an optimal method for computing the nearest orthogonal matrix [1]. There is an implementation of this algorithm available in Scipy [2].

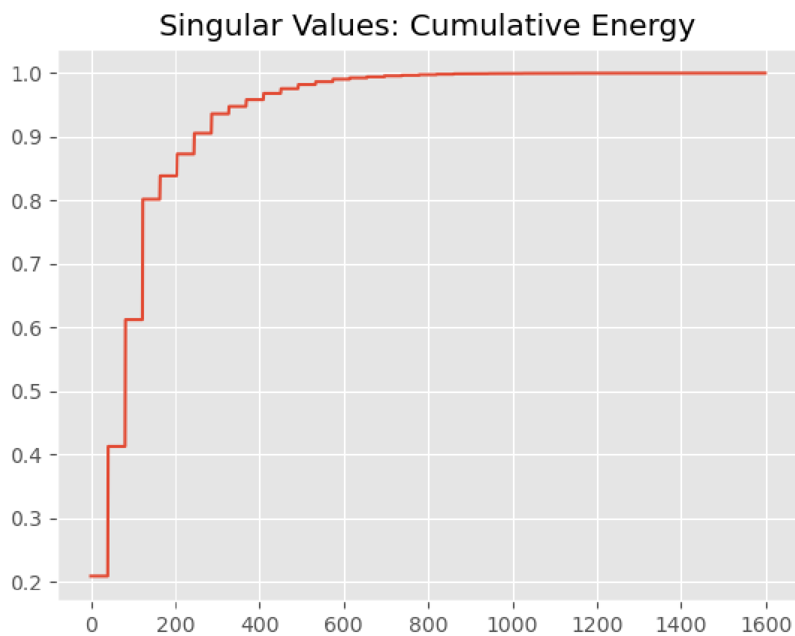

Figure 2: Singular Value Cumulative Sum

- DLO Manipulation Milestones: I am hesitant to move forward the DLO project after VPQEKF project for multiple reasons. 1) It is too advanced of a problem and I may not be able to finish it the way I want by the end of my Ph.D. studies, 2) the market oppertunity is little for such technology and at this point it is of extreme strategic importance to me to achieve low cost mass production with priority with items with the largest margin. We can keep working on the proposal but I would not spend my own time and money on it at this point. 3) To achieve DLO manipulation, first I must achieve manipulation of rigid objects first and the best approach for that in my opinion is easily available dataset. I think it is strategically important to streamline the data acquisition and annotation process. So I have shared my ideas about making a 3D object scanner and I have some ideas for generating groundtruth automatically or with minimal input. For known objects, I will define grasping normals and surfaces that can be stored onboard. Thus, the grasping problem will be reduced to pose estimation of known objects.
- 3D Scanner: It is needed for object manipulation and perception tasks.
- Pose Estimation (DLO-01): On-going under VPQEKF.
- Semantic segmentation (DLO-02): Per my discussion with Dr. Gans, I will explore DL methods for the depth or scale problem.
- Grasping Project (DLO-03): I am making this a part of the DLO project.
- PyTorch Tutorials: Transfer learning.

4 Intermediate Goals - Fall 2021:

- QEKF: Finish paper.
- UR5e: Do the tutorials.

References

[1] W. Kabsch, "A discussion of the solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 34, no. 5, pp. 827–828, 1978.

[2] "scipy.spatial.transform.rotation.align_vectors — scipy v1.8.0 manual." `https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.transform.Rotation.align_vectors.html`. (Accessed on 04/18/2022).