Final Project for CS 372

*Riley Kopp*

*Jonathan McKee*

# Algorithm, Application, Language Choice

* *Direct Linear Transform (Homography Warping)*
* *Image Stitching*
* *Python 3.6.8*

# Where It Is Used

*Homography warping is accomplished by taking matching points from two images and creating a homography matrix. Image warping works by creating a 3x3 matrix that allows any point from the first image to be translated to the correct (x, y) coordinates in the second image. The DLT algorithm is one of main methods of generating this matrix.*

## Other applications

* *Calibrating camera*
* *Placing videos or still images on a green screen.*
* *Implanting an image on a surface inside of another image*
* *Creating a 3d image representation from 2d images.*

## Alternative algorithms

* *Conic Correlation*
* *Line Correlation*

*All 3 algorithms are used to create the 3x3 homography matrix to warp the first image to be drawn into the second image. In addition, all 3 algorithms benefit greatly from normalizing the data before calculating the matrix. Normalizing the data before calculation allows the homographic algorithms to be more tolerant of noise and shifts in the planes. It is possible to use more than the minimum number of pairs, but the improvement to accuracy is gained by eliminating outliers instead of making the conversion process more accurate.*

*The DLT method works on corresponding points and relies on the matching of the points to be accurate. Because each correspondence generates two points of data in the homography matrix, we need at minimum four pairs of corresponding points.*

*The Line method on the other hand uses the cross product of two lines to determine the transformation matrix. This means we need a minimum of two pairs of corresponding points. It is also possible if we do not have the points we need in the image to generate some of the missing data using the line equations for the two matching lines.*

*The conic method uses conic shapes to match features between the objects. The conic descriptors allow us five degrees of a freedom, so we need at least two pairs of conics to achieve our eight degrees to fill in the homography matrix.*

## Reason for choice

Image stitching is a popular technology that is becoming more common place with everyday technology. Most newer cellphones now have the option to create panoramic photos in real-time which requires expanding the image stitching process to run multiple times in quick succession. There are multiple ways to accomplish image stitching, but to create a panoramic photo in real-time at high resolution requires improving the algorithm’s speed and accuracy to allow quicker matching.

One of the first steps of picking the algorithm we wanted was looking at the simplicity of the algorithm. Between the DLT style algorithms and other algorithms using different coordinate systems, the DLT style fit better with our understanding and allowed us to avoid additional post processing needed when switching coordinate systems.

After that we had to decide on what pre-processing style we wanted to use. OpenCV provides several different matchers and feature detectors each with their up and downsides. The Harris corner detector is simple, but lacks robustness when dealing with rotation, scale, and other types of transformations. SIFT is robust when it comes to dealing with invariants but SIFT normally takes a longer time to run and is proprietary. This left our main decision between the ORB detector and the AKAZE detector. The ORB detector was OpenCV’s answer to the propriety SIFT detector and the AKAZE detector was the FAST improvement to the KAZE detector. Because of that we opted to try out the newer AKAZE detector.

The next decision was to decide the max number of matches before moving into the RANSAC loop. The smaller the cap we picked, the faster the algorithm would run, but with the potential for a poor homography matrix. Therefore, we decided a 90% cutoff range would be good for both the nearest neighbor cutoff and the RANSAC would have a max run of 10,000 in case the homography matrix never creates a perfect result. If the 10,000 iterations are ran, the best homography matrix is returned from the RANSAC function. Because of the primary way to test a homography matrix while using RANSAC is via visual confirmation, we implemented a combine and draw function so we could visually confirm that the two images lined up correctly.

With the main algorithm choices existing between points, lines, and conics we chose to use the point matching. The point matching allowed us the easiest way to verify if the calculate homography function was working correctly as we could easily verify if the point had translated correctly from one image to the other. It also allowed us to generate a point map where we could connect the points with lines and quickly visually confirm correct point matches.

# How Your Project Works

* Find unique features in image 1

We use the AKAZE feature detector built into Open CV for python. The AKAZE feature detector is a multiscale feature detector that is rotation and scale invariant. To use the feature detector, we first load the image file into an array. Because of the way opencv loads the file we then need to convert the image from Blue, Green, Red into Red, Green, Blue. We also then create a grayscale copy of the image to increase the speed of processing.

* Find unique features in image 2

We then repeat the entire process with image 2.

* Find matches between features lists

Next, we use the brute force matcher to check each of the features found in steps 1 and 2 for matches. This is the brute force matcher because each feature in list 1 is checked against each feature in step 2 and the features with the highest similarity percent are selected as the match.

* Limit number of matches using nearest neighbor search

We use a nearest neighbor search to reduce the total number of matches. The nearest neighbor search checks the distance between matches for the first and second highest match for a pair of features and then divides them to create a ratio. We can then use this ratio and exclude any points with a ratio over a selected threshold. The theory behind this search is that if our feature has a high match percent to two or more features, then we have no way to determine which match is the correct one, so we exclude that feature all together. If the ratio is low, it means that the chance that we have the wrong match is low and we can then keep that match.

* (optional) Display the matching features

We have the option at this point to draw the first 50 highest ratio matches with a connecting line to allow a visual confirmation of a match.

* Use RANSAC to determine first homography matrix over threshold

RANSAC uses a random sample of 4 matches to calculate a homography matrix. We then compare the correctness of this matrix by applying it to all the points in image list 1 and then checking if they are within a tolerance of their matching point in image list 2. Once all the points are checked we divide the correct points by the total number of points to get our population fit score. The method of RANSAC we implemented exits the RANSAC loop if it gets a perfect point match or until the iterations equal the MAX\_RANSAC cutoff.

* + Calculate the homography matrix

The homography matrix is created by crafting an 8x8 matrix by layering the points of each match 2 rows at a time. Once all the rows are stacked together the matrix is ran through numpy’s SVD function in order to decompose the matrix into 3 smaller matrices (U, S, V). The V matrix is a 3x3 matrix that contains our homography matrix. But to make it work we normalize it by dividing all elements by the last element in the matrix.

* Combine and draw matched images

The first part of this step is preparing a new array to place the image into. We use the sizes of the original color images to estimate the size of the new image. We then copy image 2 directly into the new array. Next we multiply the the (X, Y, 1) coordinates of every pixel in image 1 by the homography matrix and place the pixel value into the calculated location in the new image. Once the entire image 1 has been processed we crop the new image by looking for the outer borders of the image that all contain 0 (black). The last step of the programis to output the combined image so the user can now see the two images combined into 1.

## Correctness (team only)

*[Formally, justify why your algorithm is correct or why it works. You must formally use loop invariant, pre-and post conditions, etc. Any and all properties for a given algorithm, if applicable, are required. Assume you are speaking to someone who has completed this course]*

Ratio test:

for m, n in nn\_matches:

if m.distance < (1 - CUTOFF) \* n.distance:

matched1.append(kpts1[m.queryIdx].pt)

matched2.append(kpts2[m.trainIdx].pt)

filtered.append([m])

1. Base case value
   1. nn\_matches contains all found matches between image 1 and image 2.
2. Step case
   1. 0 to m only contains matches over our cutoff value.
3. Final case
   1. All matches remaining in nn\_matches have passed the ratio test and their distance is above our cutoff.

RANSAC:

while iteration < MAX\_RANSAC and bestfit != 1.0:

points = np.random.permutation(range(0, len(match1)))[:4]

left\_points = [match1[points[0]], match1[points[1]], match1[points[2]], match1[points[3]]]

right\_points = [match2[points[0]], match2[points[1]], match2[points[2]],match2[points[3]]]

h\_mat = calc\_h(left\_points, right\_points)

fit\_pop = 0

for i, \_ in enumerate(match1, start=0):

coord = np.asarray([match1[i][0], match1[i][1], 1])

expected = [match2[i][0], match2[i][1], 1]

new\_coord = np.matmul(h\_mat, coord.T)

new\_coord = new\_coord / new\_coord[2]

outcome = (new\_coord[0] - expected[0], new\_coord[1] - expected[1])

if (-1 < outcome[0] < 1) and (-1 < outcome[1] < 1):

fit\_pop += 1

fitness = fit\_pop / len(match1)

iteration += 1

if fitness > bestfit:

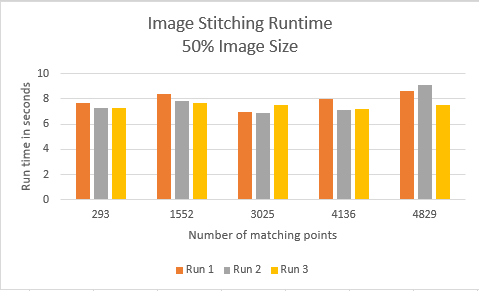
best\_h = h\_mat

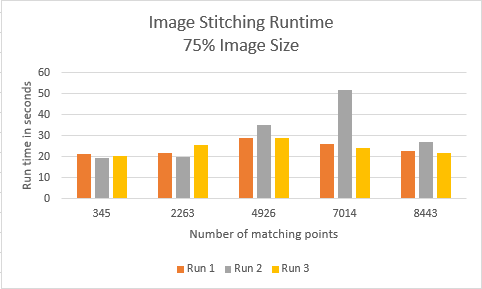
bestfit = fitness

1. Base case value
   1. Best\_h holds the best homography matrix, and bestfit has the highest fitness population.
2. Step case
   1. best\_h currently holds the best homography matrix and bestfit holds the value for the best population fit for 0 to the current iteration.
3. Final case
   1. Best\_h holds the best homography matrix for iterations 0 to iteration and is returned out of the function.

# Run time

*[Name and explain why the project has this big-O run time as the theoretical run time* **for your implementation***]*





(team only): Formally prove the run time through instruction counting, probability, or recursion analysis depending on your problem. You **will need** to use pseudocode to prove this.

# Built-in Code Correctness Tests

## Test 1; calc\_h calculation test

Test four known matching points from the S1 and S2 image sets and ensure the calculated results match the hand calculated values.

### Input

match1l = [(1092.183837890625, 853.7752685546875), (660.028076171875, 758.4589233398438), (1281.1529541015625, 343.7947082519531), (532.8238525390625, 688.1666259765625)]

match1r = [(1883.947265625, 839.6972045898438), (1452.0703125, 744.1082153320312), (2077.43603515625, 325.6315612792969), (1326.53662109375, 674.3931274414062)]

### Output

ans1 = np.reshape([.0947890104, -.00000347519559, 808.252761, -.0158891884, .0983631042, -.767619865, -.0000210778695, .00000182966520, 1.00000000], (3,3))

## Test 2; calc\_h calculation test

Test four known matching points from the S1 and S2 image sets and ensure the calculated results match the hand calculated values.

### Input

match2l = [(410.07818603515625, 635.8060913085938), (647.446044921875, 566.640380859375), (782.3014526367188, 419.570556640625), (1132.2030029296875, 1010.1979370117188)]

match2r = [(1206.5606689453125, 623.0596923828125), (1440.1629638671875, 553.3386840820312), (1574.3477783203125, 406.0274658203125), (1924.3804931640625, 996.9850463867188)]

### Output

ans2 = np.reshape([ .937084236, -.00770557503, 812.934354, -.0190458023, .974520368, 3.90652702, -.0000239529658, -.00000312357237, 1.00000000], (3,3))

## Test 3; calc\_h calculation test

Test four known matching points from the S1 and S2 image sets and ensure the calculated results match the hand calculated values.

### Input

match3l = [(1226.177001953125, 383.9251708984375), (781.3949584960938, 443.0159606933594), (1123.8994140625, 991.2728881835938), (618.127685546875, 871.2716064453125)]

match3r = [(2021.3006591796875, 366.52471923828125), (1573.3128662109375, 429.4444885253906), (1916.2684326171875, 977.9464111328125), (1410.8824462890625, 856.591796875)]

### Output

ans3 = np.reshape([ .949341489, .00427252148, 806.071557, -.0162162397, .985995762, -1.12087307, -.0000210159026, .00000329789714, 1.00000000], (3,3))

## Test 4; calc\_h calculation test

Test four known matching points from the S1 and S2 image sets and ensure the calculated results match the hand calculated values.

### Input

match4l = [(981.9266357421875, 841.3993530273438), (780.8423461914062, 448.7953186035156), (1191.7843017578125, 741.1649169921875), (1279.4105224609375, 451.46331787109375)]

match4r = [(1772.854248046875, 827.2769165039062), (1572.8653564453125, 435.0625915527344), (1984.958984375, 726.6543579101562), (2075.173583984375, 434.2452697753906)]

### Output

ans4 = np.reshape([ .950818945, .00126976874, 806.873364, -.0156713234, .986049943, -1.59223174, -.0000202314399, .00000264205453, 1.00000000], (3,3))

## Test 5; calc\_h calculation test

Test four known matching points from the S1 and S2 image sets and ensure the calculated results match the hand calculated values.

### Input

match5l = [(716.2276000976562, 452.69146728515625), (1101.726806640625, 993.9178466796875), (831.8926391601562, 303.9850769042969), (1340.210693359375, 283.180908203125)]

match5r = [(1508.5675048828125, 439.4599609375), (1893.6612548828125, 980.3692016601562), (1624.4183349609375, 290.0299987792969), (2138.21923828125, 263.7933349609375)]

### Output

ans5 = np.reshape([ .951226482, -.00223909176, 806.969830, -.0158094636, .983001045, -.422786277, -.0000199926634, .000000417320974, 1.00000000e+00], (3,3))

## Test 6; calc\_h calculation test

Test four known matching points from the S1 and S2 image sets and ensure the calculated results match the hand calculated values.

### Input

match6l = [(366.706298828125, 891.0513305664062), (1182.1109619140625, 736.4375610351562), (1013.3613891601562, 544.14501953125), (627.0460205078125, 954.373779296875)]

match6r = [(1164.088623046875, 875.904052734375), (1975.1905517578125, 721.8318481445312), (1805.2923583984375, 529.0923461914062), (1419.36083984375, 939.374267578125)]

### Output

ans6 = np.reshape([ .940287558, -.000601820060, 810.933285, -.0185216745, .981280285, 1.64164768, -.0000234528413, .00000108872770, 1.00000000], (3,3))

## Test 7; Final image check

Visually inspect the combined line from the stitched together image to ensure it is stitched on the right location.

### Input

### Output



# Program usage or README

This program has two run modes:

Python3 runStitch.py

Python3 runStitch.py [image1] [image2]

If the program detects an invalid number of arguments it will automatically default to the S1.jpg and S2.jpg sample images.

Pytest test

This command runs the built in unit tests of the calc\_h function. The test selection is needed to instruct pytest to only try and reset the built-in test for our package and not the dependencies.

# References

“Feature Matching”. (n.d.). Retrieved from https://docs.opencv.org/3.4/dc/dc3/tutorial\_py\_matcher.html.

“Singular Value Decomposition (SVD) Tutorial.” *Singular Value Decomposition (SVD) Tutorial*. Retrieved from web.mit.edu/be.400/www/SVD/Singular\_Value\_Decomposition.htm.

E. Dubrofsky and R. Woodham, Combining Line and Point Correspondences for Homography Estimation. ISVC ’08 Proceedings of the 4th International Symposium on Advances in Visual Computing, Part II, 2008. 202 -213

E. Dubrofsky. Homography estimation. PhD thesis, UNIVERSITY OF BRITISH COLUMBIA (Vancouver, 2009)

J. Kannala, M. Salo, and J. Heikkilä. Algorithms for computing a planar homography from conics in correspondence. In British Machine Vision Conference, 2006. 118

D Roosab, E Shiguemori and A Lorenac, “Comparing ORB and AKAZE for visual odometry of unmanned aerial vehicles”, <http://www.epacis.net/ccis2016/papers/paper_121.pdf>