CV Assignment 4

Kshitiz 2016051

May 2019

1 Question 1

1.1 Scientist with experience greater than 20 years

1. Andrew Zisserman : Covers the geometric principles and their algebraic representation in terms of camera projection matrices, the fundamental matrix and the trifocal tensor. (projection of data)

2. Jitendra Malik : Treated image segmentation as a graph partitioning problem and propose a novel global criterion, the normalized cut, for segmenting the graph.

3. Takeo Kanade : An iterative image registration technique with an application to stereo vision.

4. Luc Van Gool : Presented a novel scale- and rotation-invariant interest point detector and descriptor, coined SURF (Speeded Up Robust Features)

5. David Lowe : Distinctive image features from scale-invariant keypoints (SIFT features), calculated using different gaussian pyramid levels.

6. Cordelia Schmid : Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories

7. John F. Canny : Algorithm for Canny edge detector.

8. Chris Harris : Harris corner detector, a robust corner detection technique.

9. Paul V. C. Hough : Hough transform used for line, circle and polygon detection in images.

1.2 Young Researchers with experience 10-20 years

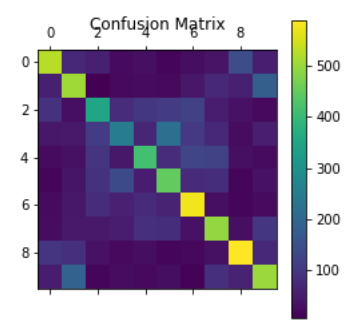
1. Muhd Amrullah : Developed a system over outdoor advertisements for facial recognition.

2. Serge Belongie : Proposed the Shape context concept : Feature descriptor for object recognition.
3. Andrew Fitzgibbon : Works on image-based rendering using machine learning.
4. Micheal Jones : Recent development of Viola-Jones face detection method.
5. Jana Košecká : Research on self-driving cars and applications of computer vision in the field.
6. Paul Viola : Recent development of Viola-Jones face detection method.
7. Nikos Paragios : Proposed geodesic active contours alongwith level sets for tracking and detection of objects in motion
8. Gang Hua : Microsoft's Facial Recognition technologies' contributor.
9. Jian Sun : Deep residual learning for image recognition. The method includes tasks of ImageNet detection and localization etc.
10. Sven Kreiss : Discovered a new particle in the search for the Standard Model Higgs boson using an ATLAS detector at LHC.

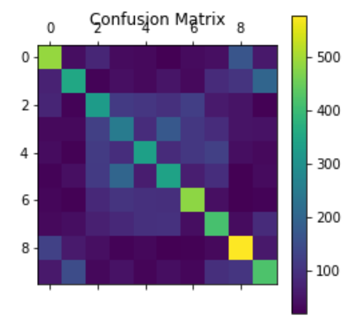
2 Question 2

2.1 Question 2a

- CIFAR data is extracted using pickle library, and was reshaped to rgb image (32x32x3), which will be used as a input to AlexNet.
- Pretrained model of alexnet was retirived and transformation is composed as following :
    1. Conversion to PIL image
    2. Resizing to (224x224) image
    3. Transforming to tensor
    4. Normalizing with mean = [0.485, 0.465, 0.406] and standard deviation = [0.229, 0.224, 0.225]
- For transforming training dataset 200 batches were created with 250 images each. Each batch is converted to tensors and then passed to AlexNet and then output is retrieved back to numpy array. This numpy array is then added to the new training data. Similar steps were followed for the testing set.
- The constructed dataset is trained on Logistic Regression classifier and accuracies are noted.
- Accuracy = 46.66%, and confusion matrix :

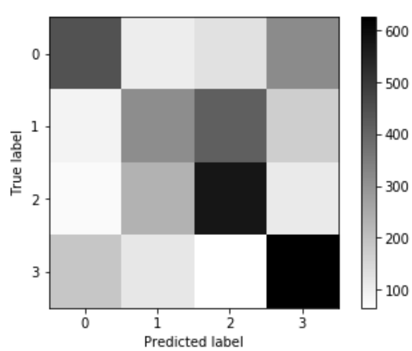


- The constructed dataset is trained on Neural Network with layers (500,500) as a classifier and accuracies are noted.
- Accuracy = 39.96%, and confusion matrix :



2.2 Question 2b

- A CNN class is written in the same architecture mentioned in the assignment.
- CIFAR data is extracted using pickle and its labels are change to 0,1,2,4 from 1,3,5,9 respectively and data related to only this labels are attached. Also the image is reshaped into 3x32x32 format for CNN.
- Data is then loaded in batch size of 20 and are passed to the CNN, with 10 epochs and 0.001 as the learing rate. And results are noted.
- Accuracy = 49% ,and confusion matrix :



3 Question 3

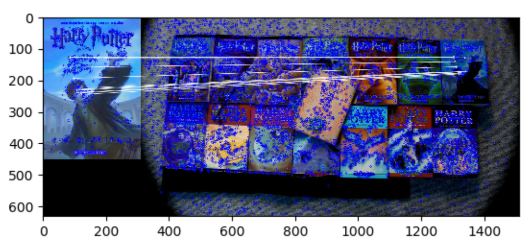
3.1 Key point matching

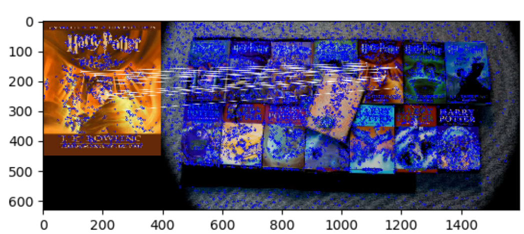
1. Inbuilt library ($sift = cv2.xfeatures2d.SIFT\_create()$) of OpenCV is used for detecting SIFT key points in images.

2. Brute force matcher ($cv2.BFMatcher()$)of OpenCV is used for matching SIFT key points in images, using knn match criteria.

3. Only major matches are saved, which were selected based on the distance measure between matching keypoints.

4. All the detected and matched key points are plotting using OpenCV library($cv2.drawMatchesKnn()$). and the results are as follows where all SIFT keypoints are colored with blue and matched keypoints are with red.





3.2 Homography

Affine matrix is constructed using the matched keypoints obtained above. For Affine transformation we need atleast 3 matches and matrix is obtained as below:

$$\underbrace{\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ & & & \vdots & & \\ x_n & y_n & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_n & y_n & 1 \end{bmatrix}}_{\substack{\mathbf{A} \\ 2n \times 6}} \underbrace{\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}}_{\substack{\mathbf{t} \\ 6 \times 1}} = \underbrace{\begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ \vdots \\ x'_n \\ y'_n \end{bmatrix}}_{\substack{\mathbf{b} \\ 2n \times 1}}$$

Now the pseudo inverse of the matrix is calculated and Affine parameters are calculated.

$$\mathbf{t} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

Note : Above pictures are from course slides. Following are the transformation matrices generated using matched keypoints.

- First image :

$$\begin{pmatrix} 4.55805589e-01 & 1.18657756e-02 & 6.65802237e+01 \\ 3.17926642e-02 & 4.57861215e-01 & 9.48988190e+02 \\ 0.00000000e+00 & 0.00000000e+00 & 1.00000000e+00 \end{pmatrix}$$

- Second image :

$$\begin{pmatrix} 6.19796136e-01 & 1.01394385e-02 & 3.58979790e+01 \\ 1.10775916e-02 & 6.36194358e-01 & 6.04682849e+02 \\ 0.00000000e+00 & 0.00000000e+00 & 1.00000000e+00 \end{pmatrix}$$

3.3 Boundary

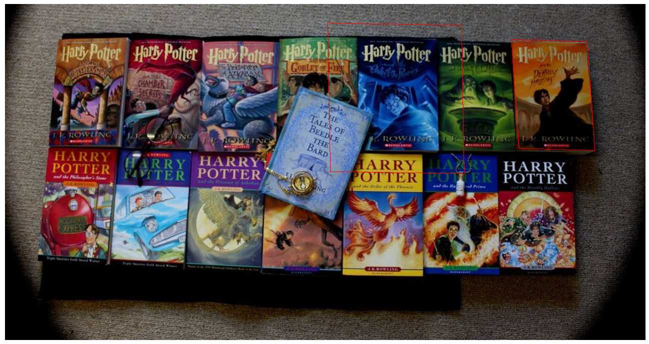
Steps:

1. Since one of the images is a subset of the other image, the subset is image is transformed on the bigger image.

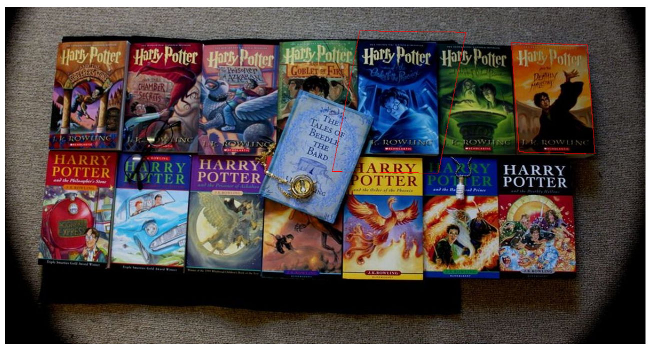
2. Transformation of image1 can be done by multiplying the pixel co-ordinates of the image1 to the homography matrix obtained above.

3. Only the border on the subset image is transformed and the pixel co-ordinate to which it maps is assigned with red color.

    - Using threshold as 0.3



- Using threshold as 0.37



4 Question 4

4.1 Key point matching

Steps:

1. Inbuilt library ($sift = cv2.xfeatures2d.SIFT\_create()$) of OpenCV is used for detecting SIFT key points in images.
2. Brute force matcher ($cv2.BFMatcher()$)of OpenCV is used for matching SIFT key points in images, using knn match criteria.
3. Only major matches are saved, which were selected based on the distance measure between matching keypoints.
4. All the detected and matched key points are plotting using OpenCV library($cv2.drawMatchesKnn()$).

4.2 Homography

Affine matrix is constructed using the matched keypoints obtained above. For Affine transformation we need atleast 3 matches and matrix is obtained as below:

$$\underbrace{\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ & & & \vdots & & \\ x_n & y_n & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_n & y_n & 1 \end{bmatrix}}_{\substack{\mathbf{A} \\ 2n \times 6}} \underbrace{\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}}_{\substack{\mathbf{t} \\ 6 \times 1}} = \underbrace{\begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ \vdots \\ x'_n \\ y'_n \end{bmatrix}}_{\substack{\mathbf{b} \\ 2n \times 1}}$$

Now the pseudo inverse of the matrix is calculated and Affine parameters are calculated.

$$\mathbf{t} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

Note : Above pictures are from course slides.

4.3 Stitching

Steps :

1. A new blank image was created and the target image is placed at the center of that blank image.
2. Size of blank image is set (not hard-coded) as when the transformed images is imposed on the new image, it does not go out the frame.
3. For each pixel co-ordinate we get a vector of size 3 [x y 1], and the pixel value is attached to co-ordinate obtained when multiplied with homography matrix.
4. 2 types of blending is done on the assignment.
   - When the final image is obtained interpolation of final image is done, so as when small image is mapped with a bigger size images, it will not leave gaps in between.
   - Average blending is done on the overlapping region.
5. Results of Average blending is shown below :