Name: Madhav

Roll no: 200657

ADOBE SCAN

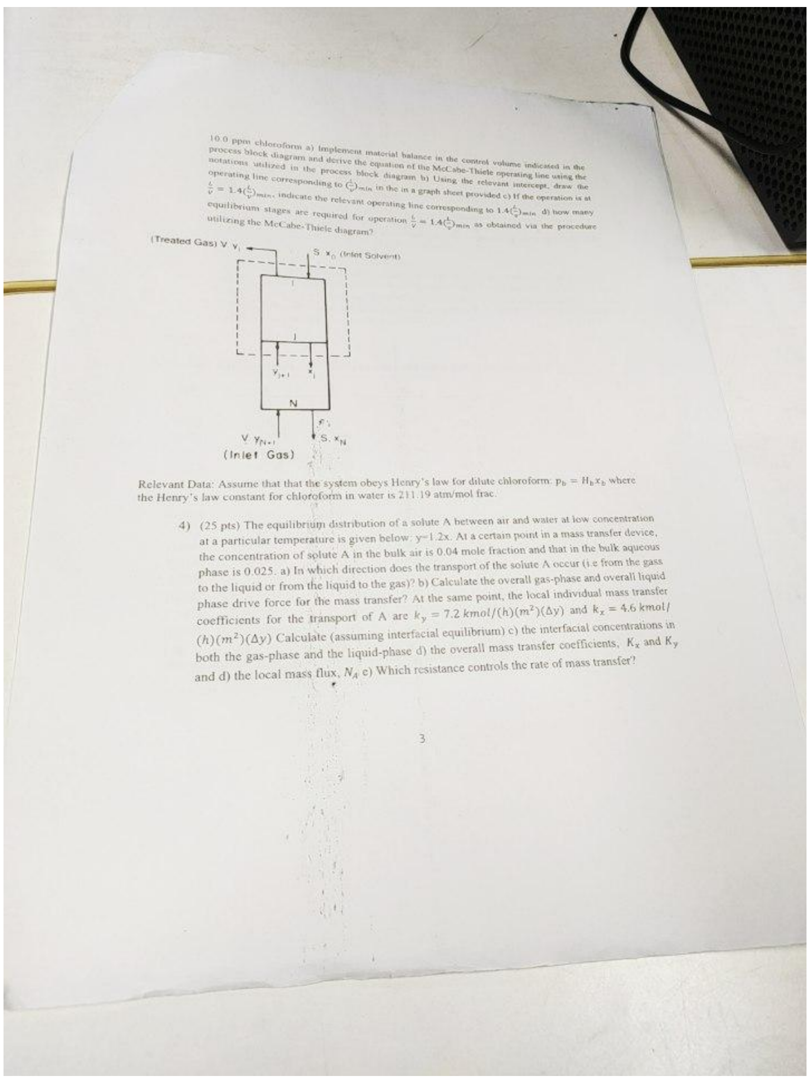
AUTOMATIC PERSPECTIVE DISTORTION RECTIFICATION

App Link

https://play.google.com/store/apps/details?id=com.adobe.scan.android&hl=en_IN&gl=US

Adobe Document Scanner - The Adobe Scan scanner app turns your device into a powerful portable scanner that recognizes text automatically (OCR) and allows you to save to multiple file formats including PDF and JPEG.

The most intelligent scanner app. Scan anything — receipts, notes, documents, photos, business cards, whiteboards — with the text you can reuse from each PDF and photo scan.



10.0 ppm chloroform a) Implement material balance in the control volume indicated in the process block diagram and derive the equation of the McCabe-Thiele operating line using the notations utilized in the process block diagram b) Using the relevant intercept, draw the operating line corresponding to $(\frac{L}{V})_{min}$ in the in a graph sheet provided c) If the operation is at $\frac{L}{V} = 1.4(\frac{L}{V})_{min}$, indicate the relevant operating line corresponding to $1.4(\frac{L}{V})_{min}$ d) how many equilibrium stages are required for operation $\frac{L}{V} = 1.4(\frac{L}{V})_{min}$ as obtained via the procedure utilizing the McCabe-Thiele diagram?

Relevant Data: Assume that that the system obeys Henry's law for dilute chloroform: $p_b = H_b x_b$ where the Henry's law constant for chloroform in water is 211.19 atm/mol frac.

4) (25 pts) The equilibrium distribution of a solute A between air and water at low concentration at a particular temperature is given below: y=1.2x. At a certain point in a mass transfer device, the concentration of solute A in the bulk air is 0.04 mole fraction and that in the bulk aqueous phase is 0.025. a) In which direction does the transport of the solute A occur (i.e from the gass to the liquid or from the liquid to the gas)? b) Calculate the overall gas-phase and overall liquid phase drive force for the mass transfer? At the same point, the local individual mass transfer coefficients for the transport of A are $k_y = 7.2\ kmol/(h)(m^2)(\Delta y)$ and $k_x = 4.6\ kmol/(h)(m^2)(\Delta y)$ Calculate (assuming interfacial equilibrium) c) the interfacial concentrations in both the gas-phase and the liquid-phase d) the overall mass transfer coefficients, K_x and K_y and d) the local mass flux, N_A e) Which resistance controls the rate of mass transfer?

3

Fig. (1): Input Image (captured by mobile camera)

STEP-2: (Pre-processing step for perspective transformation)

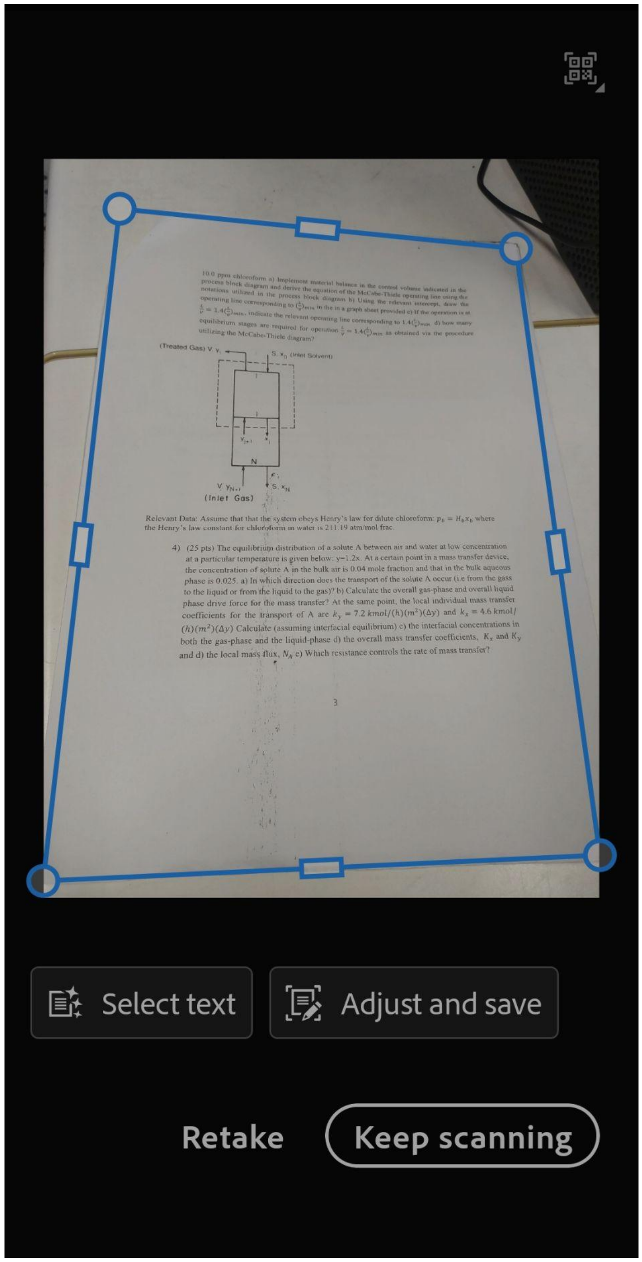


Fig. (2): Edge & Corner of document detection

Image Processing algorithms used: Edge detection (using Sobel or Canny), Dilation and Erosion, Harris corner detection

OpenCV functions:

cv2.canny(), cv2.dilate(), cv2.erode(), cv2.cornerHarris(), cv2.circle(), cv2.line()

Edge detection:

Convert the image into grayscale to perform the following edge detection algorithms

Sobel operators method -

Using Sobel operators to find gradients in both the x and y directions separately using the following two Sobel operators.

$$\mathrm{G_x} = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathrm{A} \quad \text{and} \quad \mathrm{G_y} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * \mathrm{A}$$

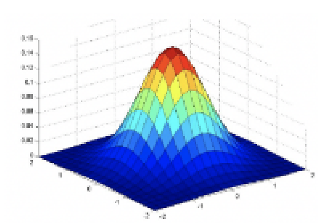
Then find the gradient and its direction as

$$\mathbf{G} = \sqrt{\mathbf{G}_x{}^2 + \mathbf{G}_y{}^2}$$
$$\mathbf{\Theta} = \operatorname{atan2}(\mathbf{G}_y, \mathbf{G}_x),$$

Canny algorithm -

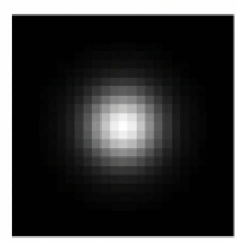
This algorithm consists of 4 steps - Smooth derivatives o Thresholding o Thinning o Linking

Reduce noise and smoothen derivatives by convolving the image with the gaussian kernel



Gaussian kernel

$$G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$



Find the magnitude and direction at each pixel

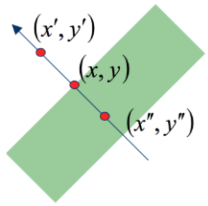
$$g_x(x,y) = \partial f_s(x,y)/\partial x \qquad g_y(x,y) = \partial f_s(x,y)/\partial y$$

$$M_s(x,y) = \|\nabla f_s(x,y)\| = \sqrt{g_x^2(x,y) + g_y^2(x,y)} \qquad \alpha(x,y) = \tan^{-1}\left[\frac{g_y(x,y)}{g_x(x,y)}\right]$$

Non-max suppression - checks whether the pixel is local maxima in grad direction

Edge occurs when the gradient reaches a maxima

Suppress non-maxima gradient even if it passes the threshold



$$M(x,y) = \begin{cases} |\nabla G|(x,y) \text{ if } |\nabla G|(x,y) > |\nabla G|(x',y') \\ \qquad \& |\nabla G|(x,y) > |\nabla G|(x'',y'') \\ 0 \qquad \text{otherwise} \end{cases}$$

x' and **x**'' are the neighbors of **x** along normal direction to an edge

Use hysteresis to link the edges on the edge map and connectivity analysis to detect edges. Avoid streaking near the threshold value.

a. find all edge points using $TH\ high$

b. from each strong point follow both side directions ⊥ to the edge normal

c. in that direction, construct the contours of connected edge points

d. mark all points greater than $THlow$

Thus from the canny or Sobel algorithm, we get an edge map

Dilation and Erosion:

Dilation and erosion techniques using the structuring elements are used to segment the image.

Assuming E to be a Euclidean space or an integer grid, A a binary image in E, and B a structuring element.

Dilation is $A \oplus B = \left\{ z \middle| \left(\hat{B}\right)_z \cap A \neq \varnothing \right\}$ and erosion is $A \ominus B = \{z | (B)_z \subseteq A\}$

The structuring element is used for probing and expanding the shapes contained in the input image.

Perform closing operation for the broken edges to get filled.

- Closing
  - useful for filling
    - small holes
    - gaps

$$A \cdot B = (A \oplus B) \ominus B$$

The segmented image is the input for corner estimation. This step of Adobe Scan can also be used for OCR (Optical Character Recognition).

Harris Corner detection:

Harris corner detector is a mathematical operator that finds a variation in the gradient of an image. It is rotation, scale, and illumination variation independent. It sweeps a window w(x,y) with displacement u in the x direction and v in the y direction and calculates the variation of intensity.

$$\mathrm{E(u,v)} = \sum_{x,y} w(x,y)[I(x+u,y+v) - \mathrm{I(x,y)}]^2$$

Maximizing the RHS gives us corners

Let's denote $\mathrm{M} = \sum_{x,y} w(x,y) \begin{bmatrix} \mathrm{Ix}^2 & IxIy \\ IxIy & Iy^2 \end{bmatrix}$

Now $\mathrm{E(u,v)} = [\mathrm{u} \;\; \mathrm{v}]\, \mathrm{M} \begin{bmatrix} u \\ v \end{bmatrix}$

A score is calculated for each window, to determine if it can possibly contain a corner

$$\mathrm{R = det(M)\text{-}k(trace(M))^2}$$

where det(M) = λ1 λ2; trace(M) = λ1+ λ2; λ1 and λ2 are the eigen values of M. R depends on eigenvalues of M. A window with a score R greater than the threshold is selected and points of local maxima of R are considered as a “corner”

The four points of interest is obtained by selecting the minimum and maximum values in the corner points array as shown in Fig. (2)

STEP-3: (perspective correction)

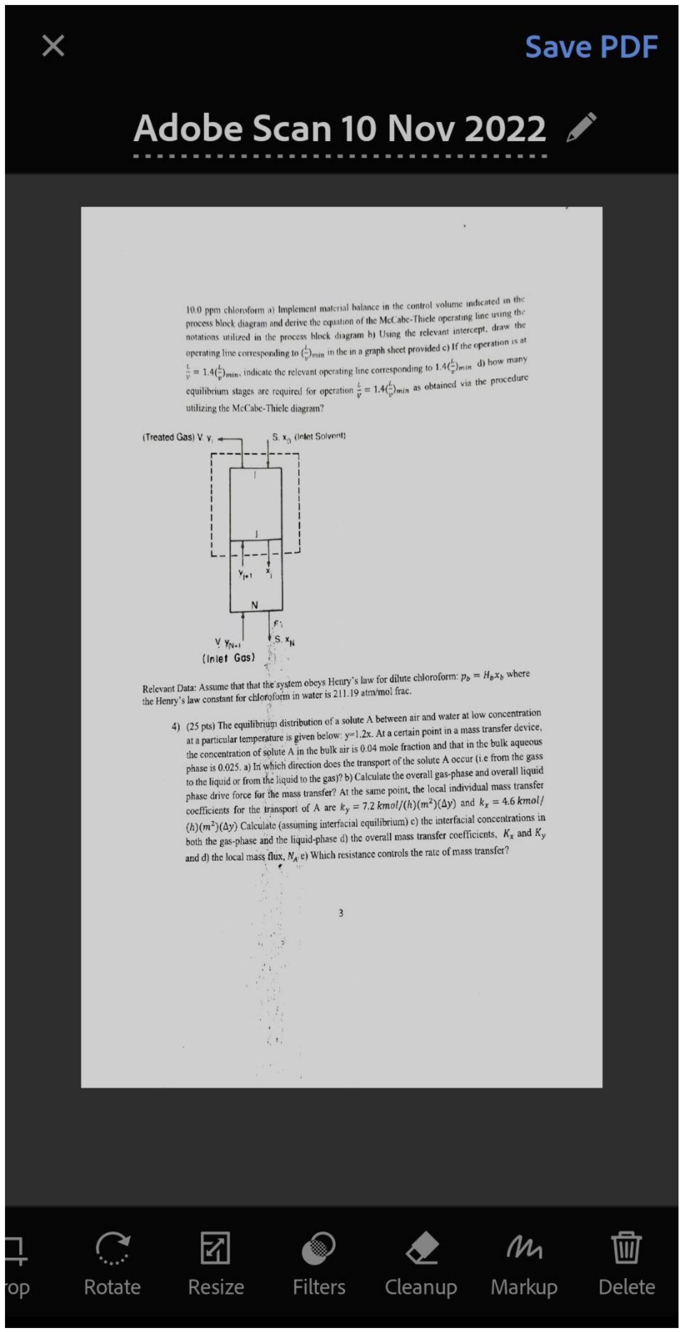


Fig. (3) Perspective-transformed Image

Image Processing algorithms used: Perspective correction using plane homography method.

OpenCV functions:

Mat	***cv::getAffineTransform*** *(const **Point2f** src[], const **Point2f** dst[])*
	Calculates an affine transform from three pairs of the corresponding points

Mat	***cv::getPerspectiveTransform*** *(**InputArray** src, **InputArray** dst)*
	Calculates a perspective transform from four pairs of the corresponding points.

void	***cv::warpAffine*** *(**InputArray** src, **OutputArray** dst, **InputArray** M, **Size** dsize, int flags=**INTER_LINEAR**, int borderMode=**BORDER_CONSTANT**, const **Scalar** &borderValue=**Scalar**())*
	Applies an affine transformation to an image.

void	***cv::warpPerspective*** *(**InputArray** src, **OutputArray** dst, **InputArray** M, **Size** dsize, int flags=**INTER_LINEAR**, int borderMode=**BORDER_CONSTANT**, const **Scalar** &borderValue=**Scalar**())*
	Applies a perspective transformation to an image.

To build a parallel projection of the image from the perspectively distorted image, the plane homography is used. The homography can be computed by knowing the relative positions of the four points on the perspectively distorted image and the positions of the transformed image to be constructed.

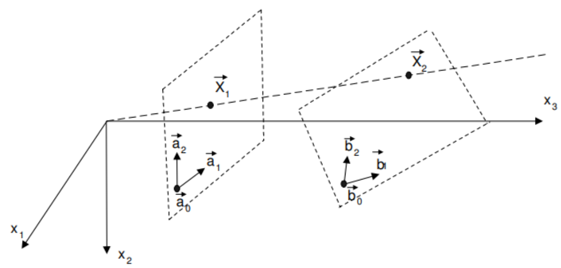


Fig. (4) Plane Homography technique

A 2D homography is defined as a 3X3 homogeneous matrix that maps any point p(x,y) on plane π to its corresponding point p′ (x′,y′) on π’. Then, these points are related by the *estimated* homography ***H***, as follows: p′ = H • p

$$\begin{pmatrix} wx' \\ wy' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Solving x′ and y′: $x' = \frac{ax+by+c}{gx+hy+i}$ and $y' = \frac{dx+ey+f}{gx+hy+i}$.

The two linear equations with respect to unknown coefficients of matrix H is as

$$ax+by+c-gxx'-hyy'-ix'=0$$
$$dx+ey+f-gxy'-hyy'-iy'=0$$

It can be generalized as

$$ax_i+by_i+c-gx_ix'_i-hy_iy'_i-ix'_i=0$$
$$dx_i+ey_i+f-gx_iy'_i-hy_iy'_i-iy'_i=0$$

Finally equation (11) can be written as

$$P' = H \cdot P \rightarrow A_i \cdot h = 0 \text{ for } i = 1,2,3,4.$$

where A_i is 2x9 matrix, h is a 9x1 matrix and the result is a 2x1 matrix. If all the four matrix equations are put into one equation, it results with the following

$$\begin{pmatrix} A1 \\ A2 \\ A3 \\ A4 \end{pmatrix} . h = 0$$

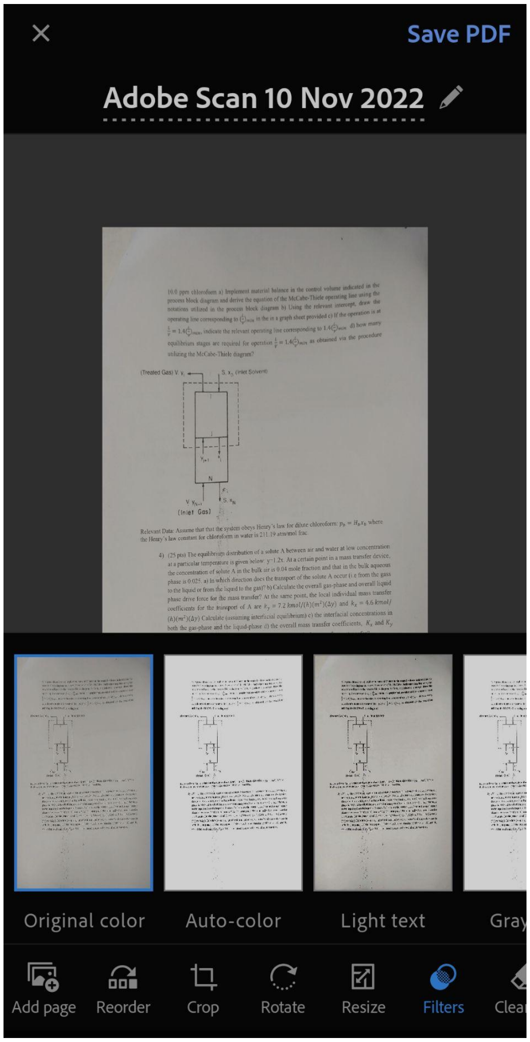
The transform is given by the solution as explained above. The elements $A_{1\text{-}8}$ can be given as

$$
\begin{bmatrix}
X1 & Y1 & 1 & 0 & 0 & 0 & -X1*x1 & -Y1*x1 \\
0 & 0 & 0 & X1 & Y1 & 1 & -X1*y1 & -Y1*y1 \\
X2 & y2 & 1 & 0 & 0 & 0 & -X2*x2 & -Y2*x2 \\
0 & 0 & 0 & X2 & Y2 & 1 & -X2*Y2 & -Y2*y2 \\
X3 & Y3 & 1 & 0 & 0 & 0 & -X3*x3 & -Y3*x3 \\
0 & 0 & 0 & X3 & Y3 & 1 & -X3*y3 & -Y3*y3 \\
X4 & Y4 & 1 & 0 & 0 & 0 & -X4*x4 & -Y4*x4 \\
0 & 0 & 0 & X4 & Y4 & 1 & -X4*y4 & -Y4*y4
\end{bmatrix}
$$

Given four corners of the screen mapping to four corners of the skewed perspective object, we have 8 equations. It implies that there are 8 linear equations and 9 unknowns. Add constraint $\|h\|=1$ which will simplify to $A \bullet h=0$ subject to $\|h\|=1$. The solution is obtained using Singular Value Decomposition for the 4-point correspondence between two planes.

When the transformation is applied to perspective images, the perspective distortion is rectified as highlighted in Fig. (3).

Adobe Scan has a magic color feature that automatically thresholds the text from its background. That automatic correction feature is shown below:



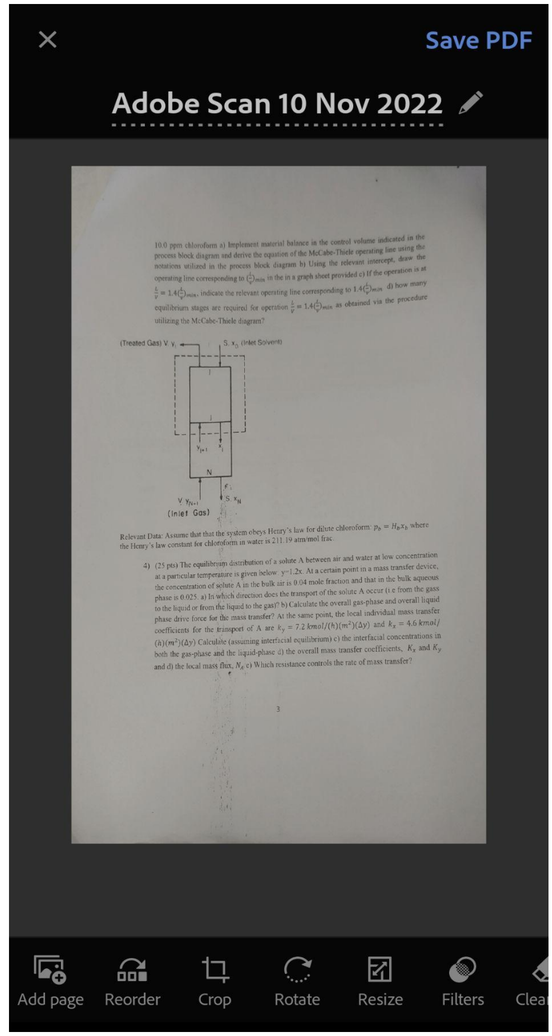


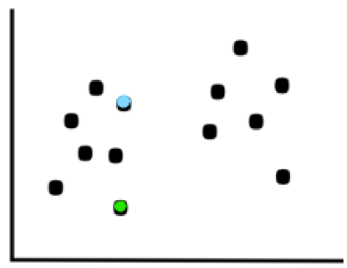
Fig. (5) Left: To get back the Original color from Auto Color,

Right: Original colored perspective corrected image

This concludes the entire algorithm of automatic perspective distortion correction of Adobe document scanner.

ANALYTICAL QUESTION

- Briefly describe k-means algorithm steps and cluster the following feature points using 2 seed points given showing each step till clusters are formed finally.



Solution:

K-means algorithm:

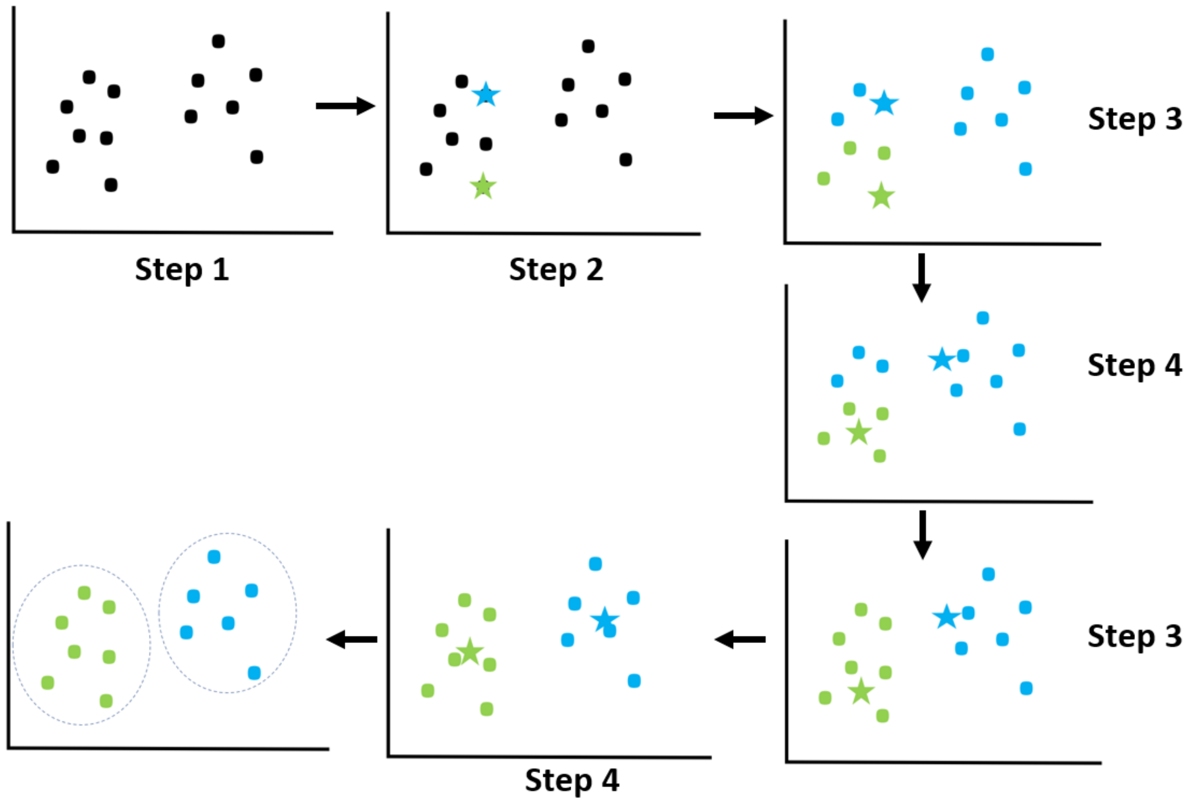
- **Input:** $x^{(1)}, x^{(2)}, ..., x^{(n)}$
- **Output:** Set of clusters $C_1, C_2, ...C_k$
- **Initialization:** Randomly pick ***k*** centroids $z^{(1)}, z^{(2)}, \ldots, z^{(k)}$
- **Itereate** until convergence or up to iterations T
  - **Assignment:** Assign each point to its closest centroid

    for each $j = 1, .., k$

    $C_j = \{i | \text{s.t. } x^{(i)} \text{ is closest to } z^{(j)}\}$

  - **Update:** Recompute centroids with newly assigned points

$$z^{(j)} = \frac{1}{|C_j|} \sum_{i \in C_j} x^{(i)}$$



Step 1
Step 2
Step 3
Step 4
Step 3
Step 4

MCQs

1. Clustering or Cluster analysis is the method of grouping entities based on similarities. Defined as an unsupervised learning problem that aims to make training data with a given set of inputs but without any target values. Choose all the correct options from the following based on different clustering methods discussed in class
    a. In Region growing, pixels in different groups may have similar properties
    b. The probability of a data point getting clustered into a single cluster increases as bandwidth increases in mean-shift clustering method
    c. Training data labels are not available in unsupervised learning
    d. The time complexity of the mean-shift algorithm is $O(n^3T)$
    e. Mean-shift is more robust to outliers than k-means
    f. The number of clusters formed in the mean-shift method can be controlled using window size (w) parameter
    g. Watershed segmentation doesn't necessarily require initial seed point(s) as input by the user

Solution: (c,e,g)

2. Choose all the correct answers from the following
    a. JPEG and PNG are image file formats and have different compression algorithms
    b. jpeg uses a sub-image size of 16x16
    c. Run length encoding of the following quantized table has the term [(5,4),13]

$$\begin{bmatrix} 100 & -60 & 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 13 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

    d. DWT is preferred over DCT in JPEG 2000 because of lesser computations and higher compression ratio
    e. Fast Fourier Transform (FFT) has a higher root-mean-square error than Discrete Cosine Transform (DCT) but is computationally faster
    f. Hoffman coding is used in communications and networks because of little loss of information while coding RLE outputs
    g. None of these

Solution: (e)