List of Figures

1.1 Word written in (a) Hand printed script. (b) Cursive script. (c) Mixed script ................................. 2
1.2 Visual appearance variation of same handwritten words written by five writers. .............................. 3
1.3 Visual appearance variation of same handwritten words written by same writer. ............................ 3
1.4 Basic steps of Hand written text to digital text(taken from IAM database). ................................. 3

2.1 Example of 62 legal words written by three writers of our database. ................................. 17
2.2 (a) Resultant vector $R'$ of gravitational force(G) and translational force(T). (b) Path described by the pen-tips under resultant force are parallel. (c) Comfortable writing: Path described by the balls under resultant force and pattern generating force are always non-intersecting if the starting point separation is greater than two times of maximum possible deviation(d). (d) Uncomfortable writing: Path described by the balls are always nonintersecting if the starting point separation is greater than twice the maximum possible deviation(d) plus deviation due to slope change(b). ....... 21
2.3 (a) Touching component of a handwritten text. (b) Bounding box with height greater than threshold-height is formed. (c) Bounding boxes with half of its previous height are formed and widths are refined. ................................. 23
2.4 Vertical overlap between of two boundary boxes: $(h_1 \cap h_2)/h_1$ . . . 24
2.5 Text having nine connecting boxes with their corresponding centroids. ........................................................................ 24
2.6 Centroids arrangement. (a) Tree for the first line identification. (b) Centroids present in the first line are connected. (c) Tree for second line identification. (d) Centroids present in the second line are connected. ............................................... 25
2.7 Results: a. and d. Segmented text lines for approximately straight handwritten document. b. Segmented text lines for downward skewed handwritten document. c. Segmented text lines for upward skewed handwritten document. .................................................. 27
2.8 Results: (a) Segmented text lines at first attempt. Errors occur in first line after ‘:’ and in second line near the dot of ‘i’ of window. (b) Corrected text lines after merging the lines whose average height difference is less than 20. .................................................. 28
2.9 Example of line segmented image of 62 legal words written by one writer. ........................................................................ 29
2.10 Example of word segmented text-line following projection method. ........................................................................ 29
2.11 a. Sample word. b. Distribution of black to white transition for each row of sample word. c. Core region estimation of word. d. Base line estimation. e. Slope corrected word. ............................. 30
2.12 (a) A sample text-word. (b) Modified word after applying transformation (Tr). (c) Word after discarding the Black pixels present in first and forth quartile of modified word. (d) Best fit straight line estimation. (e) Text word after eliminating the black pixels whose perpendicular distance from the estimated line is greater than threshold. (f) Estimation of Base line of the text word. (g) slope corrected word. .................................................. 34
2.13 Sample words and Slope corrected words. ........................................ 35
2.14 a. Slope corrected word. b. Suitable Gabor filter estimates the global slant angle. c. slant corrected word. ................................. 37
2.15  a. Slope corrected word. b. Spectrum of image word. Vertical spacial frequency is for slope. Global slant correction is done by shearing image at angle $\theta$. c. Word after global slant correction. d. Core region of word is eliminated. e. Word after local slant correction.

2.16 Length of structural element (L) selection if $\theta$ is maximum slant angle of stroke and b is stroke width.

2.17 (a) Sample text-word whose slant to be corrected. (b) Connected component present in the word. (c) Connected component is split into local strokes. (d) Slant corrected word.

2.18 (a) Sample text-word whose slant to be corrected. (b) Slant corrected word.

2.19 (a) Sample words, (b) Sample words after slope correction, (c) Sample words after slope and slant correction, (d) Slope and slant corrected sample words after space compaction where gray lines are removed.

3.1 (a) An image containing a sawtooth type shape, and (b) plot of maximum accumulator value versus angle obtained by Hough transform.

3.2 Illustrates, in raster order, 12 Arnold transform images of sawtooth where $N = 72$ and $T = 12$.

3.3 Illustrates different Arnold transformed images up to three levels (top: L=1; middle: L=2; bottom: L=3) for the image of handwritten word ‘Twenty’, where period is $T_p=12$.

3.4 Distribution of stroke orientation with a step of 15° in the range $[-90°, +90°]$ of different Arnold transformed images of ‘Twenty’ for Level-1.

3.5 Arnold Transform based directional feature for three level quadtree of three words: a. ‘one’ b. ‘Eighty’ and c. ‘Seventeen’.

3.6 Confusion matrix for ISIHWD database.

3.7 Performance (measured as % accuracy) for each class of ISIHWD database.
3.8 Plot to show the demonstration of positive correlation between test accuracy versus size of training sample in each class. .............. 63
3.9 Confusion matrix for CENPARMI database. ......................... 64
3.10 Performance (measured as % accuracy) for each class of CENPARMI database. ................................................... 65
3.11 Performance (measured as % accuracy) for each class of IAM200 data-set. ......................................................... 66
4.1 Preprocessing pipeline: Sample word → Resized word → Resized thin word. ......................................................... 68
4.2 Some black pixels with their surrounding eight pixels of a thin handwritten word. ................................................. 72
4.3 Top row: some possible three run curvatures for thin word. Bottom row: not possible three run curvatures for thin word. .......... 72
4.4 weights of the 8-neighbouring positions.* denotes the candidate pixel. ............................................................. 72
4.5 (a) Curvelet having minimum index $C_{ij} = 1 + 3 - 4 = 0$. (b) Curvelet having maximum index $C_{ij} = 14 + 5 - 4 = 15$. .......... 72
4.6 a. A preprocessed words with a portion marked by red circle. b. The intensity of pixels present in the marked area 0 for foreground and 1 for background. c. Marked area after labeling by $L_{pj}$ using equation 4.1.4, * indicates background pixel. d. Marked area after labeling by $L_c$ using equation 4.1.3, * indicates background pixel. 74
4.7 Confusion matrix for ISIHWD database. .......................... 81
4.8 Performance (measured as % accuracy) for each class of ISIHWD. ................................................................. 82
4.9 Confusion matrix for CENPARMI database. ...................... 83
4.10 Performance (measured as % accuracy) for each class of CENPARMI. ............................................................. 84
4.11 Demonstrating positive correlation between recognition accuracy and size of training sample in each class. ..................... 85
4.12 Performance (measured as % accuracy) for each class of IAM. ................................................................. 86
5.1 Input image ‘One’ of size $[81 \times 207]$ followed by 5 convolution feature maps having 5 convolution mask of size $[5 \times 9]$ and max pooling layer having filter size $2 \times 2$ with stride 2. ....................... 89
5.2 Top: Input word ‘One’. Middle row: Output of five convolution feature maps. Bottom row: Output of pooling layers. .......... 90
5.3 (a) Structure of DCNN model. ........................................ 91
5.4 Distance transformed word images of ‘One’, ‘Rupees’, ‘fifteen’ written by three different writers. .......................... 93
5.5 Training and validation error of DCNN for Word image (left), Distance transformed word image (right) on ISIHWD database (top row), CENPARMI database (middle row) and IAM200 (bottom row) (source code used from MatConvNet package [116]). .......... 96
5.6 Plot to show the demonstration of positive correlation between test accuracy versus size of training sample in each class of CENPARMI database. ........................................... 97
5.7 Confusion matrix for ISIHWD database for DCNN feature of word image. ........................................ 103
5.8 Performance (measured as % accuracy) for each class of ISIHWD database. ........................................ 104
5.9 Performance (measured as % accuracy) for each class of word image. 105
5.10 Performance (measured as % accuracy) for each class of CENPARMI database. ........................................ 106
5.11 Performance (measured as % accuracy) for each class of IAM-200 database. ........................................ 107
6.1 Schematic diagram of the proposed Word recognition system ... 109
6.2 Performance (measured as % accuracy) for each class of CENPARMI database. ........................................ 116
6.3 Confusion matrix of proposed system on CENPARMI database. ... 120
6.4 Performance (measured as % accuracy) for each class of ISIHWD database. ........................................ 121
6.5 Confusion matrix of proposed system on ISIHWD database. .... 122
6.6 Performance (measured as % accuracy) for each class of IAM200. 123