Topographic Map Processing

Arun Agarwal
Professor
University of Hyderabad

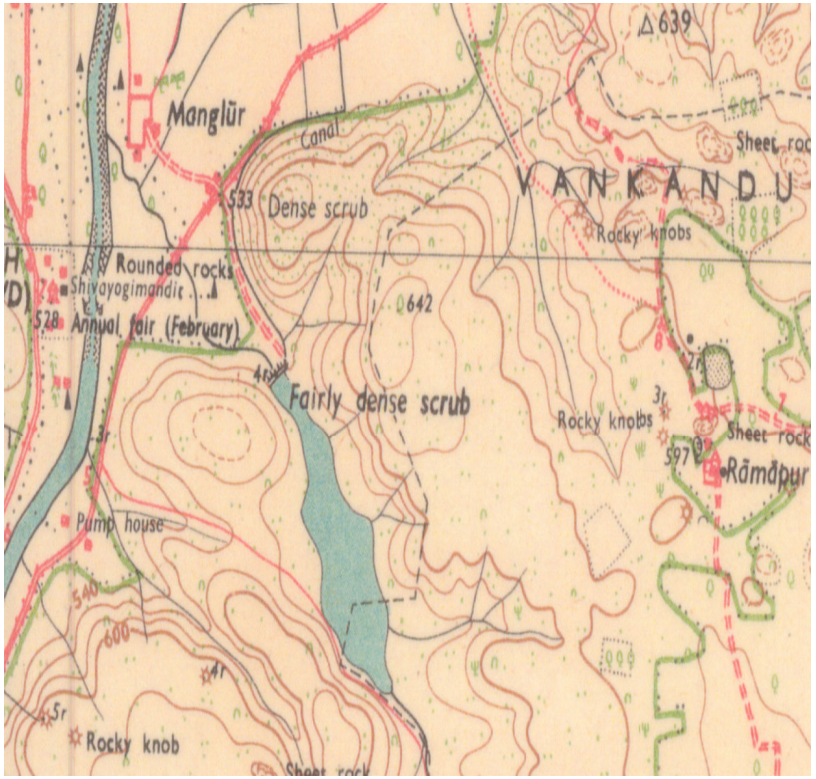
Acknowledgements

B. Sandhya
PhD Scholar
University of Hyderabad

Prof. C R Rao and Dr Rajeev Wankar
Faculty, Department of Computer and Information Sciences
University of Hyderabad

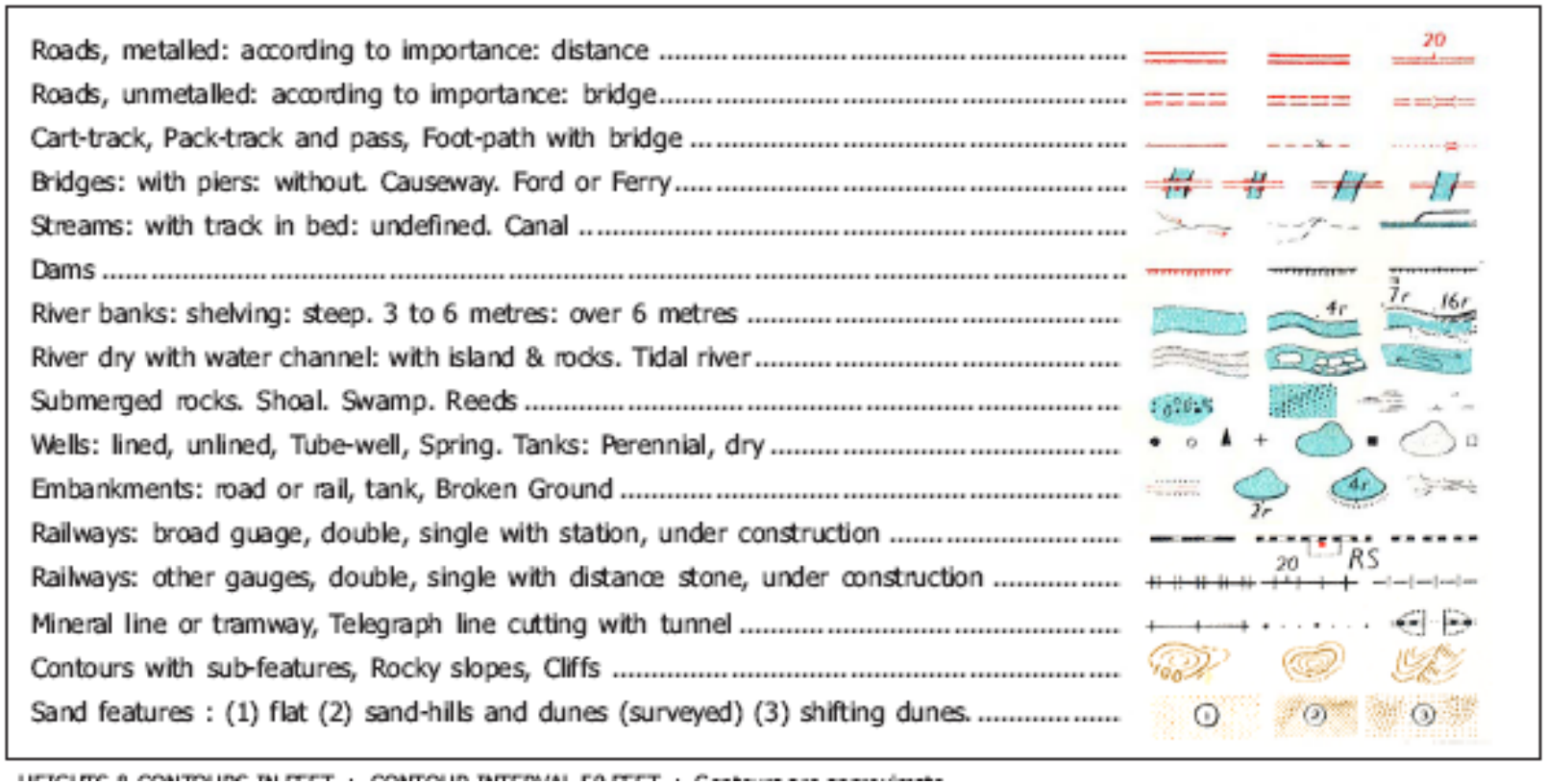
Topographic Map

- Colour line drawings which portray geographic and topographic information about the terrain.[1]
- Feature category
  - Linear features - lines : contour lines, roads etc
  - Area features - regions : vegetation, water bodies etc
- Use colour to help the user distinguish between different feature categories.
- At a typical sampling resolution of 120 pixels per centimetre, a 20 x 30 cm map would yield an image of 2400x3600 pixels.

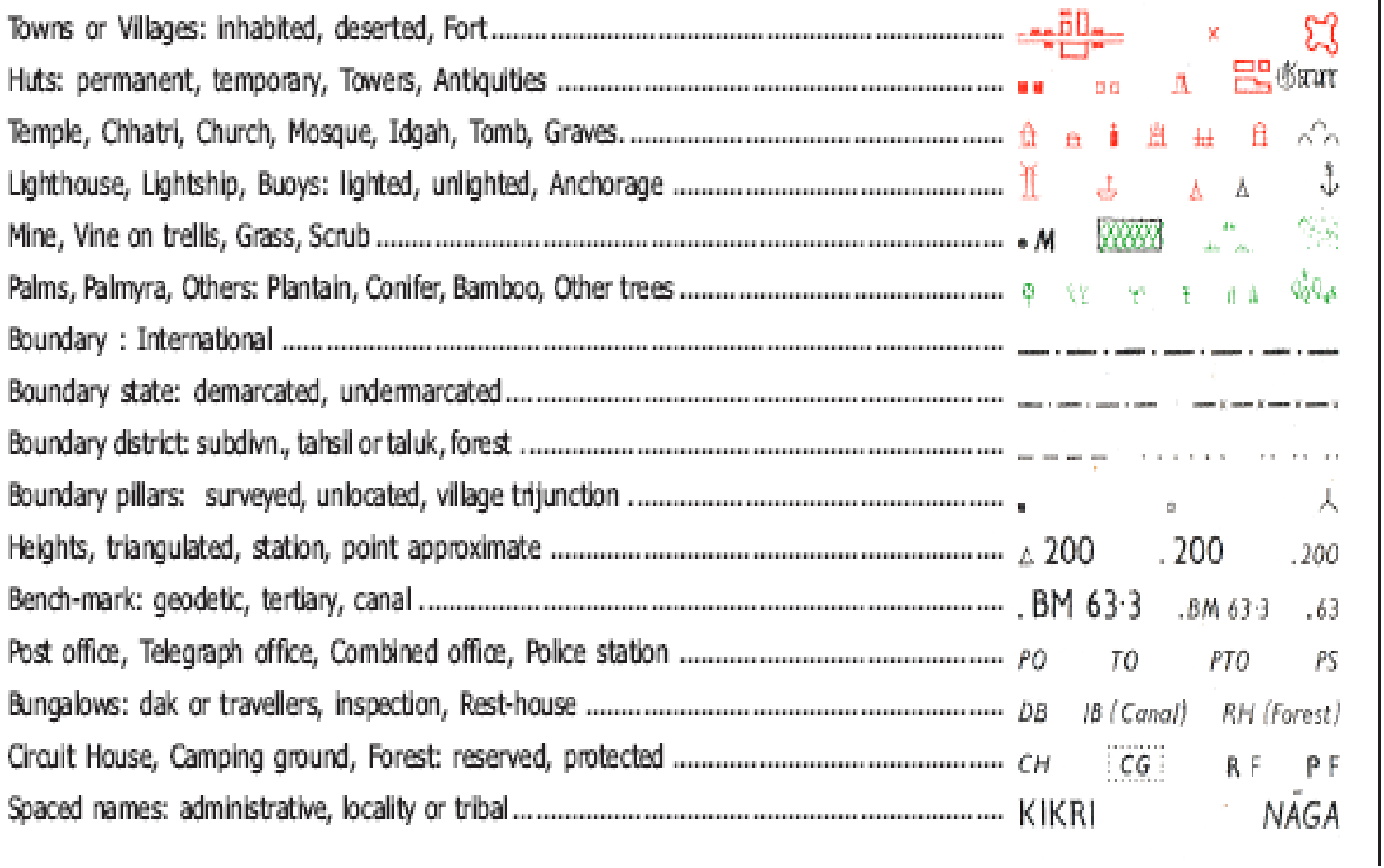


part of 1:50000, 48M/13 First Edition, Survey Of India
Scanned at 300dpi

Conventional Symbols used in Topomaps
SOI standards



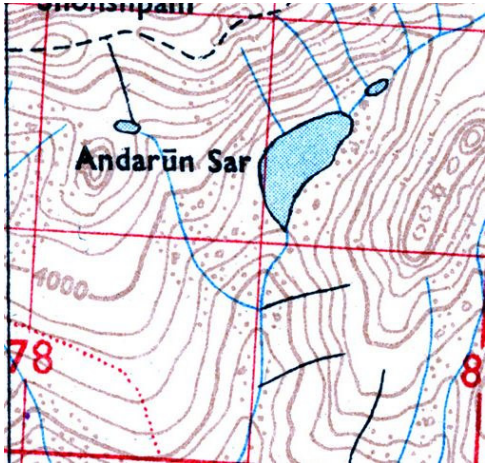
Conventional Symbols – contd..

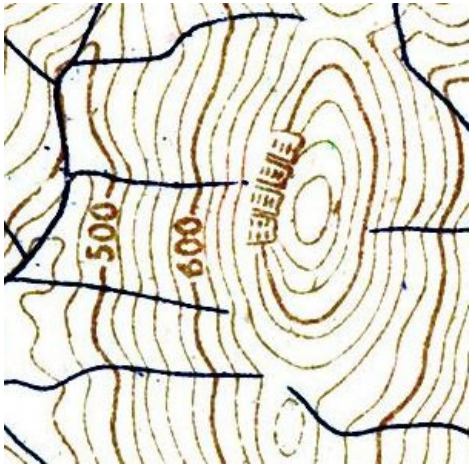


Source : Survey Of India

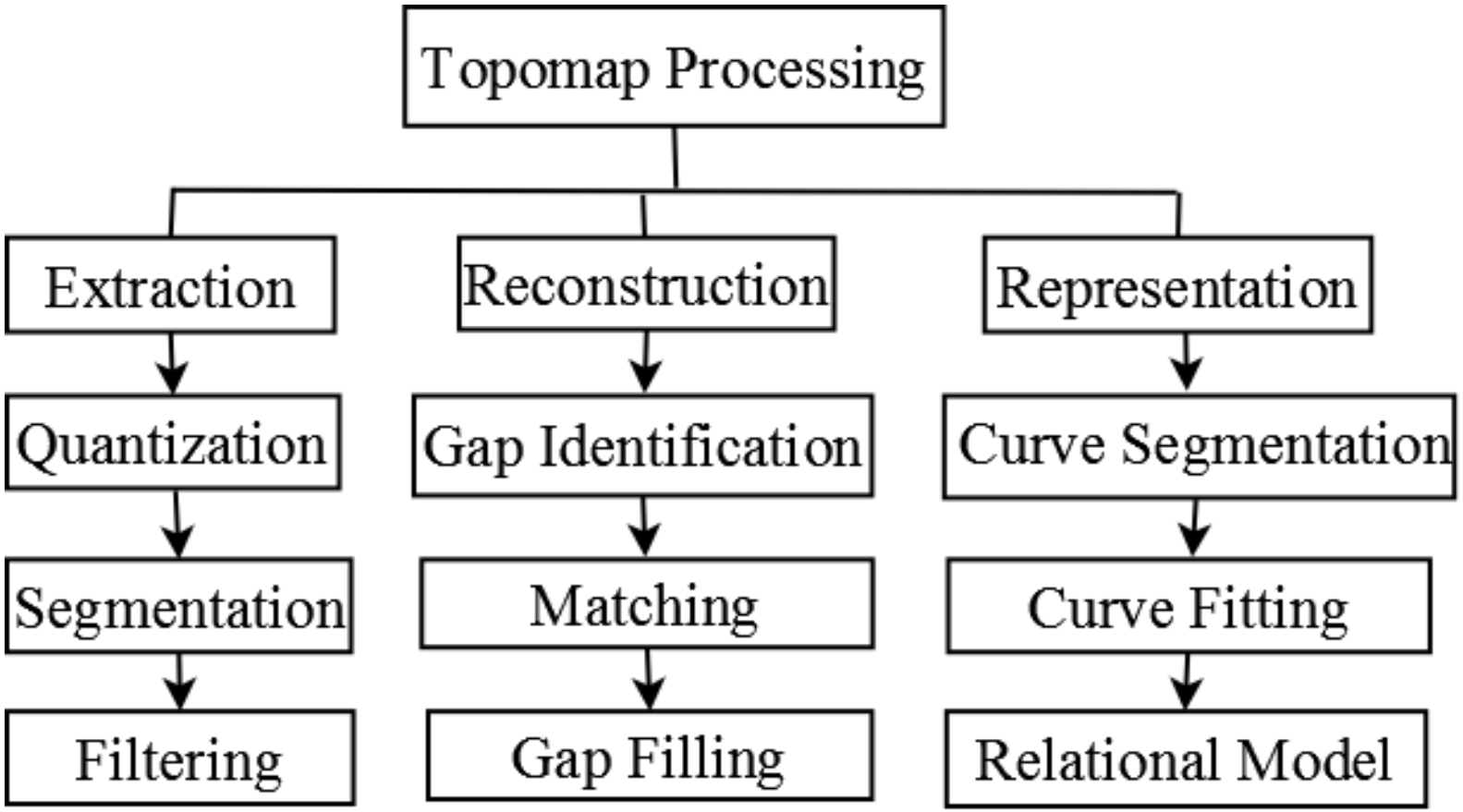
Contour Lines

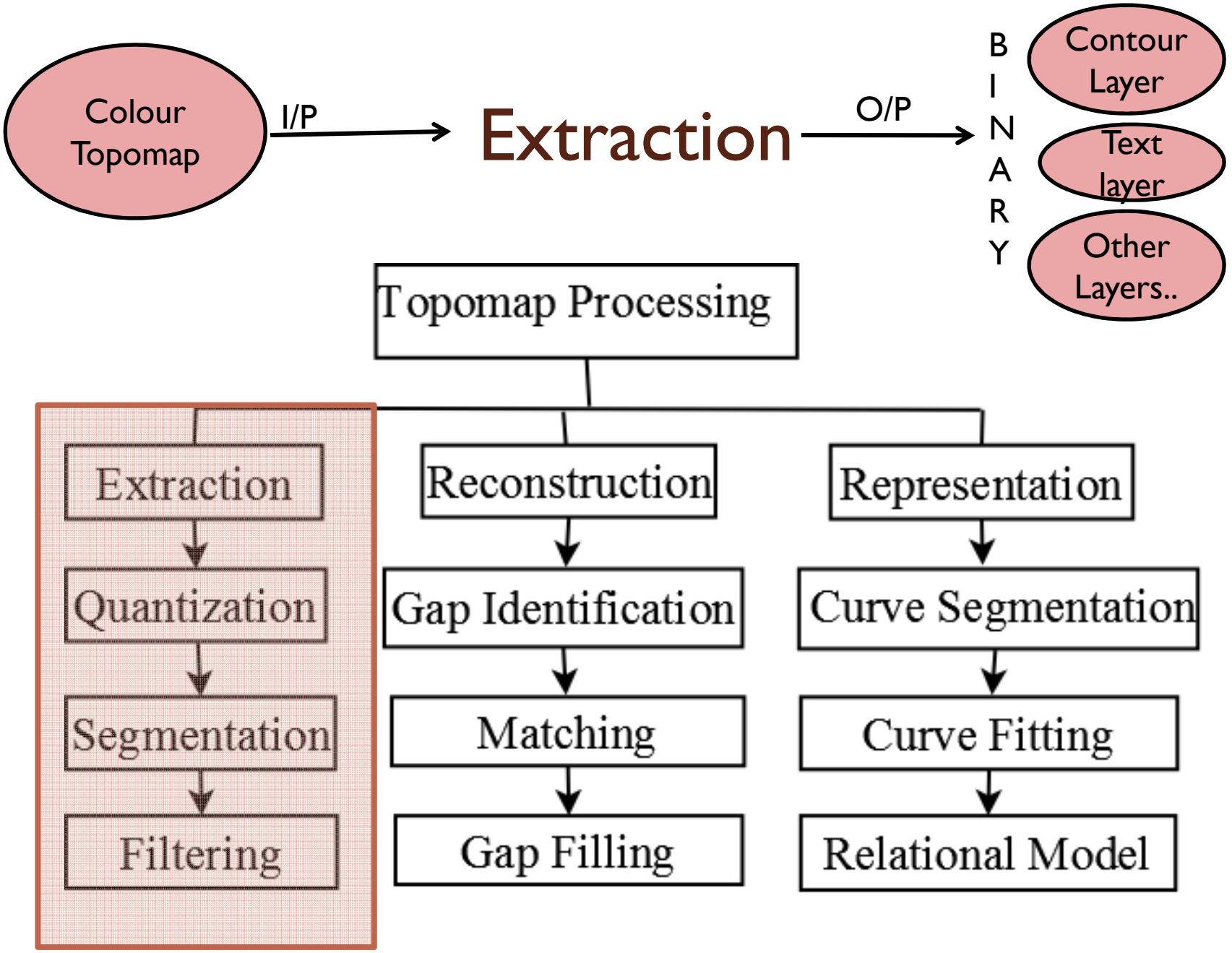
- Contour lines are curve features that connect contiguous points of the same terrain elevation (altitude).
- Convey three dimensional information of the terrain's surface.
- Contour lines are represented by lines with 2-4 pixels width when scanned at a resolution of 300dpi
- labelled indicating the elevation values above the mean sea level.





Topomap Processing: Phases




Colour Topomap
I/P
Extraction
O/P
B I N A R Y
Contour Layer
Text layer
Other Layers..
Topomap Processing
Extraction
Quantization
Segmentation
Filtering
Reconstruction
Gap Identification
Matching
Gap Filling
Representation
Curve Segmentation
Curve Fitting
Relational Model

Segmentation/Clustering

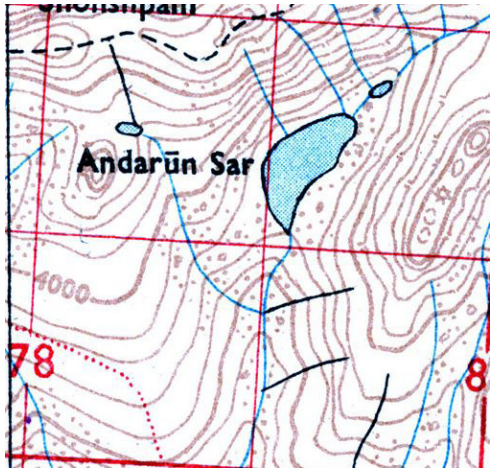
- Topomap is divided into various layers. Each layer represents one distinct feature of a map like lakes, contour lines, buildings, symbols etc.
- Main feature to distinguish various layers is color. So most of the clustering algorithms are based on color.
- Literature survey indicates that extensive work has been done in improving the segmentation of topomaps using various segmentation algorithms such as threshold based, histogram based, region based (region growing and merging), and hybrid approaches.
- Each of the segmentation algorithms use any one of the colour spaces like RGB, HIS, Lab etc. But none of them can dominate the others for all kinds of colour images. Selecting the best colour space still is one of the difficulties in colour image segmentation (CJSW01).
- Focus is to select the colour feature set, which can efficiently segment any kind of topographic map.

Clustering algorithm

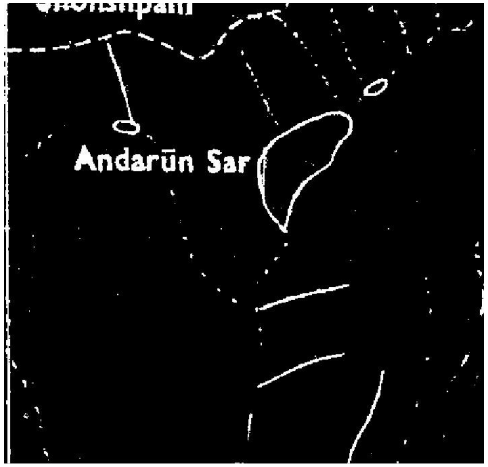
The topomap segmentation has been performed using a non parametric histogram based clustering algorithm proposed by Khotanzad et.al.

- A three dimensional histogram of the quantized image is computed by counting the number of pixels in each cell of the cube
- Clustering is performed by locating the peaks of the multidimensional histogram using peak climbing algorithm
  - The clustering operation requires the examination of all the neighbouring cells of all non-empty cells. The number of neighbours for each discrete cell is a function of the dimension of the space. The number of possible neighbours in the two and three dimensional representation are 8 and 26 respectively.
  - By examining the counts of neighbours of a particular cell, a link is established between that cell and the closest cell having the largest count in the neighbourhood. If two or more neighbouring cells have the same count for a particular cell,one of them is arbitrarily chosen as a parent.
  - At the end of the link assignment each cell is linked to one parent cell, but can be parent of more than one cell.
  - A peak is defined as being a cell with the largest density in the neighbourhood, i.e. a cell with no parent.
  - A peak and all the cells which are linked to it are taken as a distinct cluster representing a mode of the histogram.

Results on Topomap 1

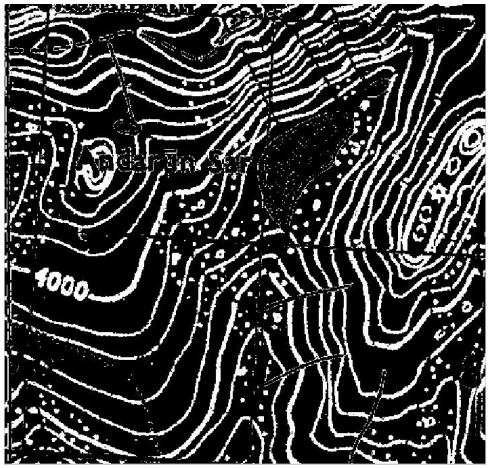


Topomap 1

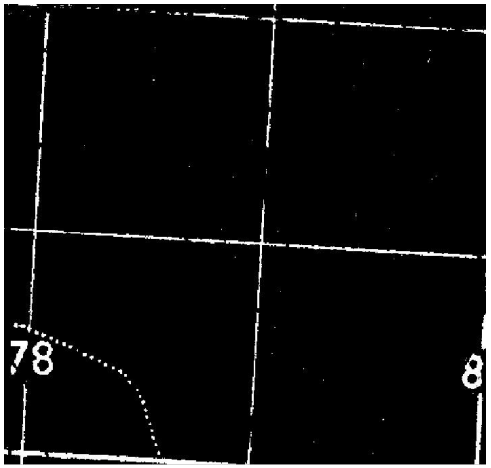


Text & Boundaries (Black)

Streams (Blue)

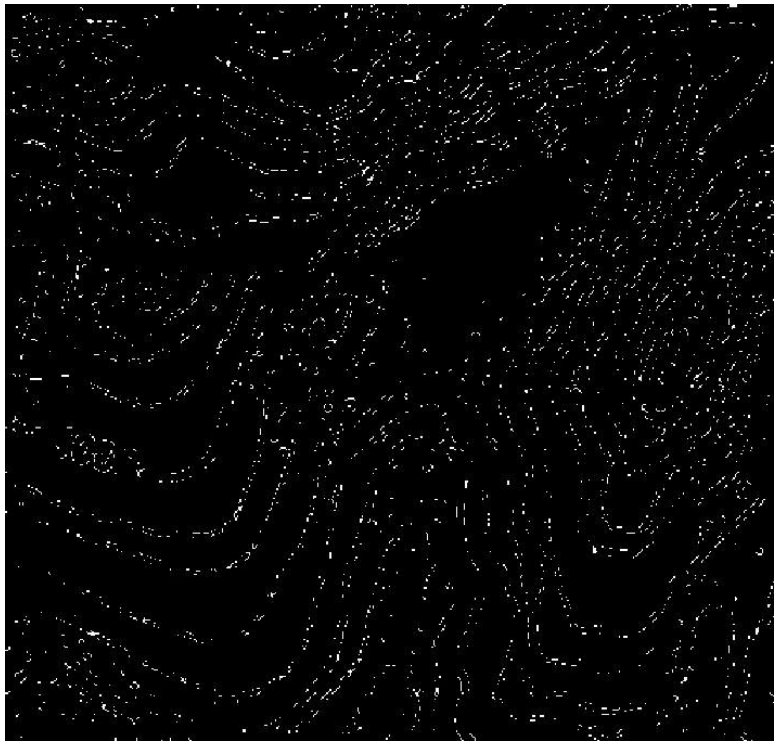


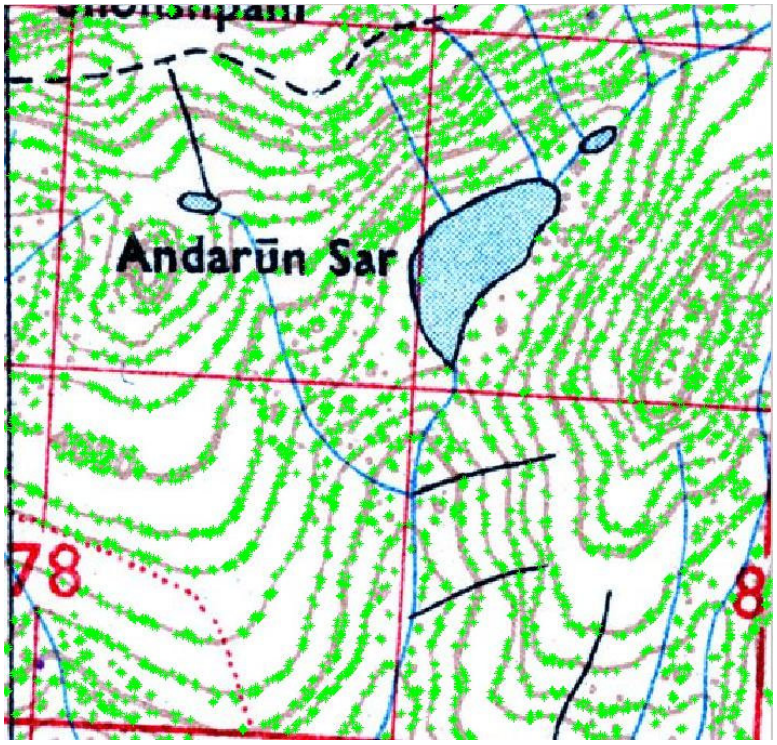
Contour layer (Brown)



Grid Lines (Red)

Difference of pixels between RGB & PCA

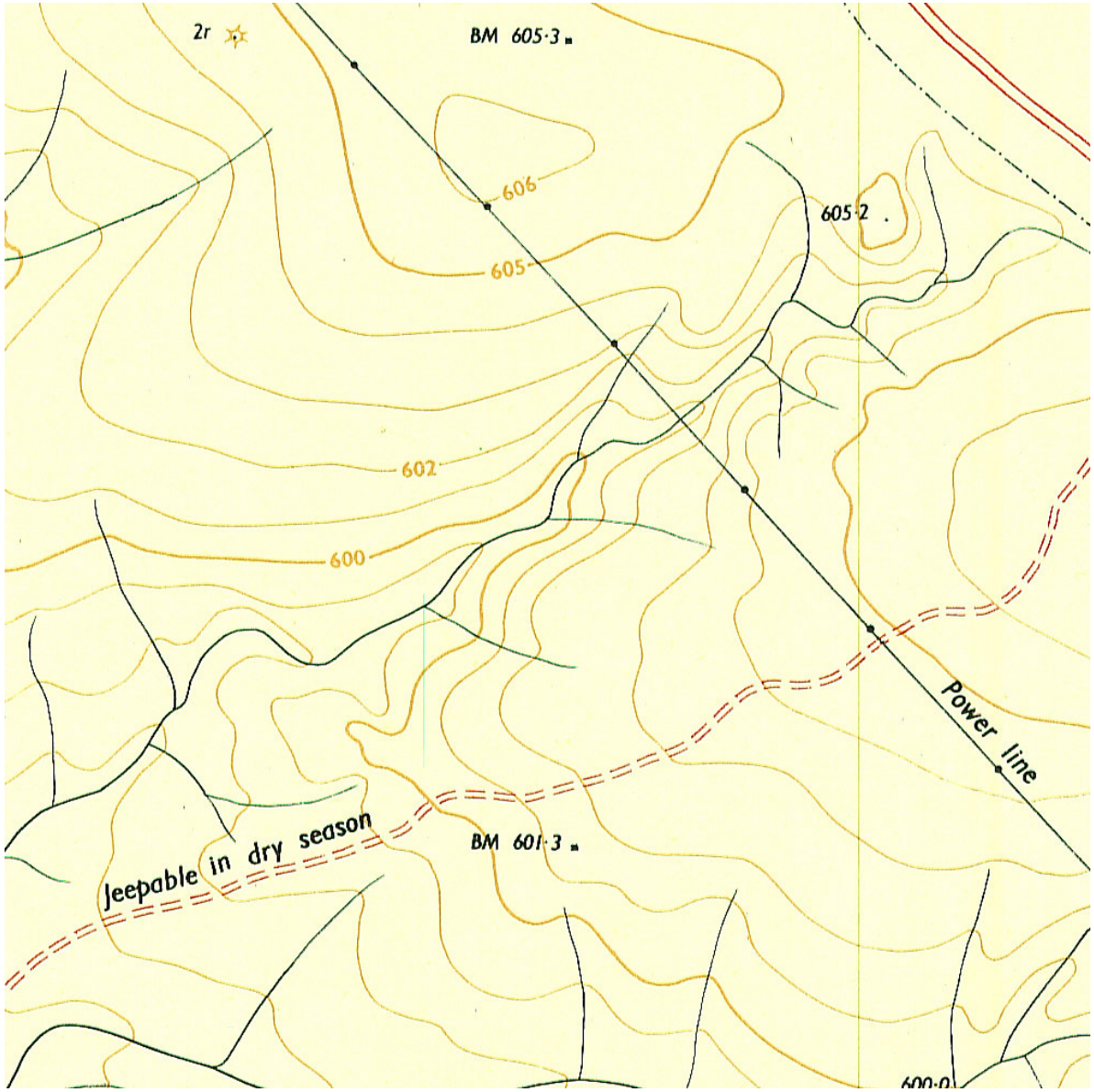




4811(about 9% of total contour pixels) pixels are more when PCA is used, which have been plotted on the original topomap (with green). As it can be observed, most of them belong to contours.

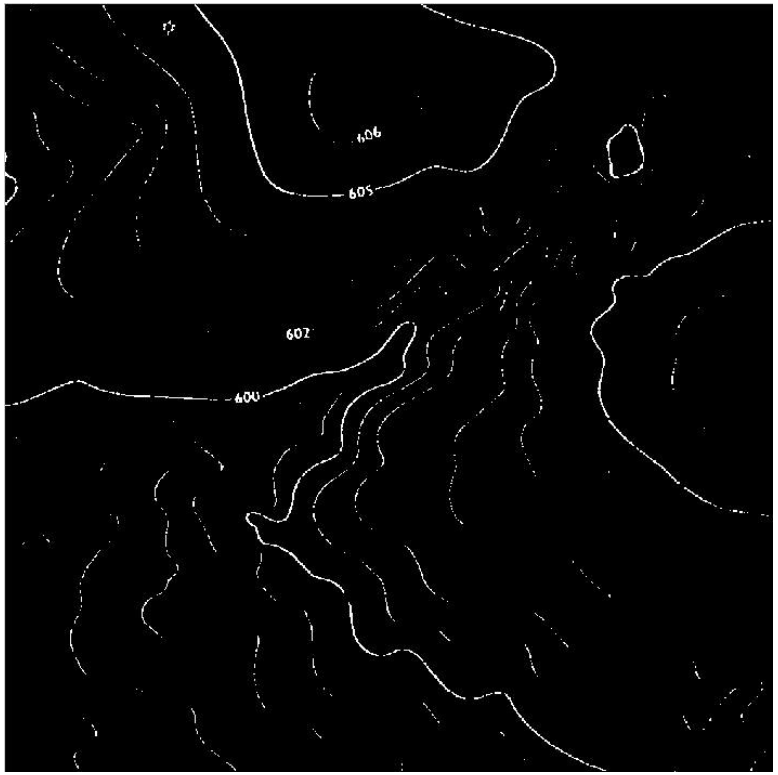
Topomap 2

Portion of UoH topomap
Scanned at CMSD
Size: 1001 x 1001



Contour layer after median filtering the cluster

Using RGB space



Using PCA transformed space

Test Images

(a) Topomap 3

(b) Topomap 4

(c) Topomap 5

(d) Topomap 6

(e) Topomap 7

(f) Topomap 8

(g) Topomap 9

(h) Topomap 10

(i) Topomap 11

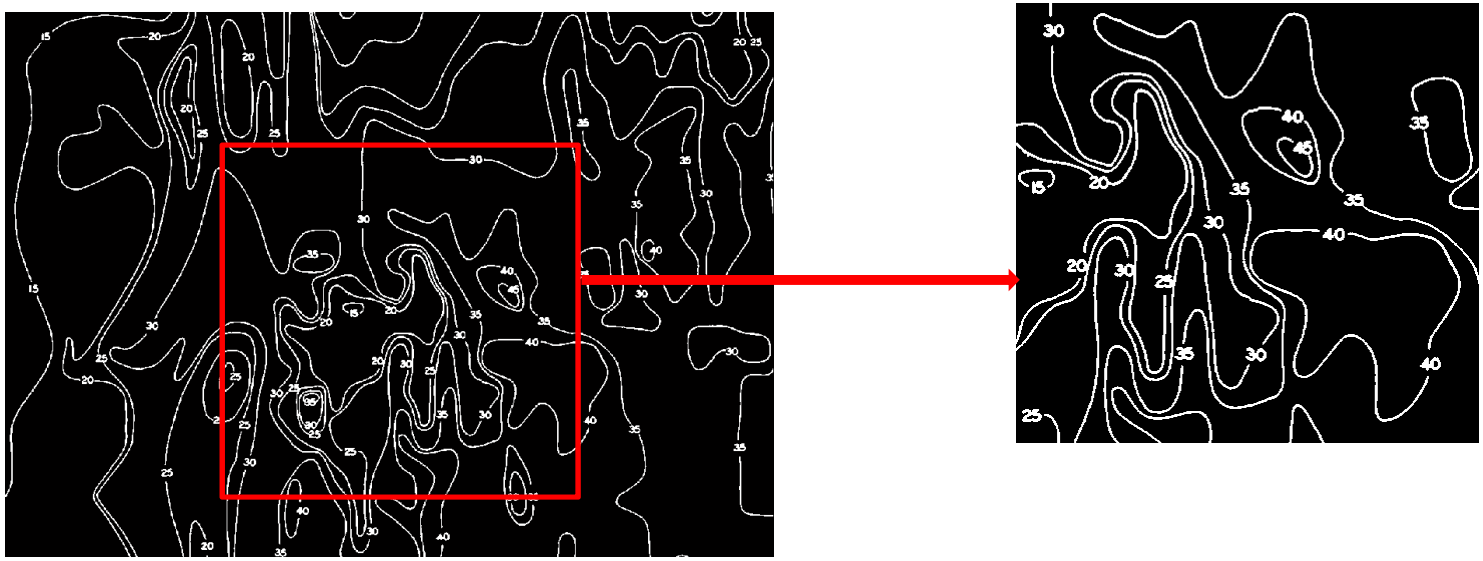
(j) Topomap 12

Filtering of Contour layer

Separation of text/graphics - Survey

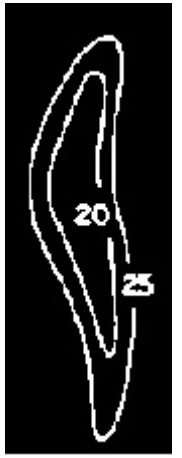
- Automatic segmentation of text/symbols from documents has been one of the fundamental aims in graphics recognition.
- One way has been to separate text from documents based on color - This is applicable when text is represented using a different color from the other graphics in the document.
- For example Joachim Pouderoux et. al [PGPG07] have extracted toponyms from map based on color.
- Second branch of research is to extract text from documents without color features. Two cases have to be considered here again
  - Text which is not connected to the graphics
  - Text which is connected/touching to the graphics.

Input Image

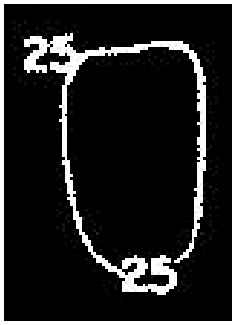


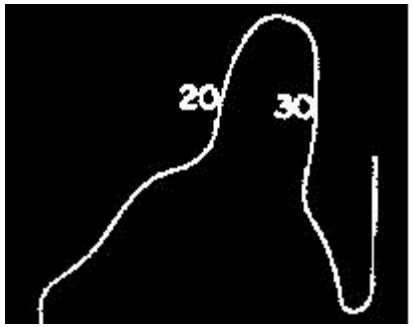
Filtering Labels – Altitude values

- Cannot be removed during Extraction stage as they are of the same color as the contour lines.
- Are of uniform width across the topographic map.
- Depending on the position of altitude values, the tags can be divided into two parts
  - Text in between two contour lines, not connected to any of the contour lines
  - Text connected to a contour line(s)



Text not connected to contour lines





Text connected to contour lines

Text Filtering – Proposed Approach

Step 1: Region Properties Extraction

Step 2: Cluster components based on properties extracted in step 1.

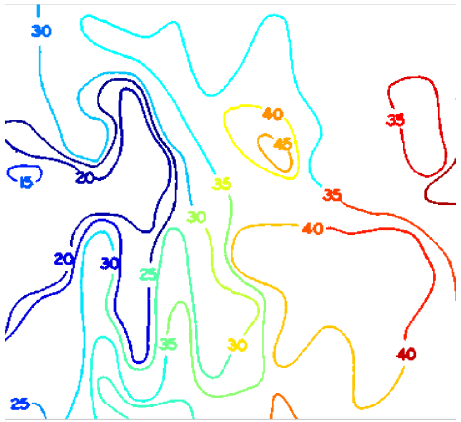
Step 3: Filter text/symbols not connected to contours

Step 4: Filter text connected to contour lines.

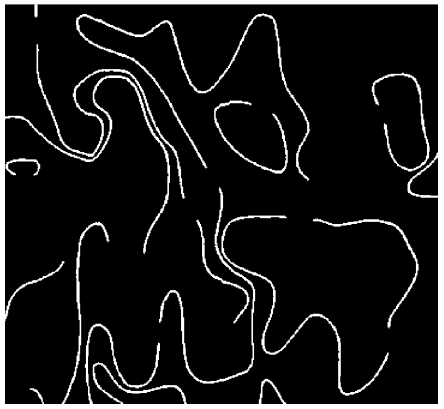
Step 5: Filter Noise

Step 1. Region Properties Extraction [Ref: KGG02]

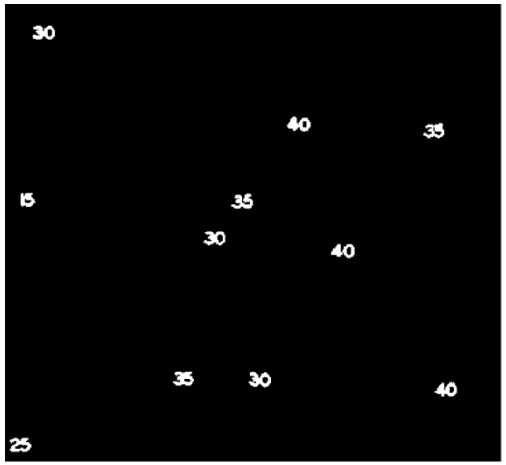
- Perform Connected Component Analysis (CCA)
- Components obtained can be classified into three categories
  - Segments of Contour lines only
  - Altitude tags
  - Contour lines with altitude tags attached



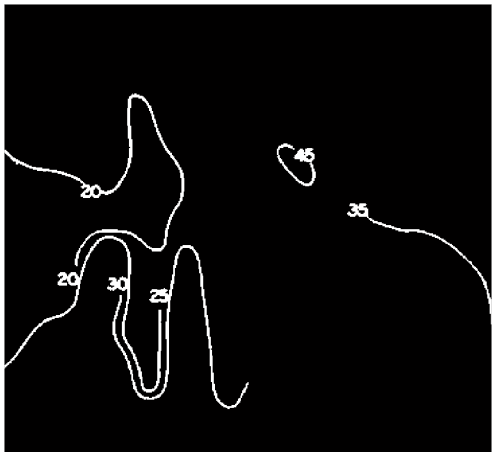
CCA



Segments of Contours

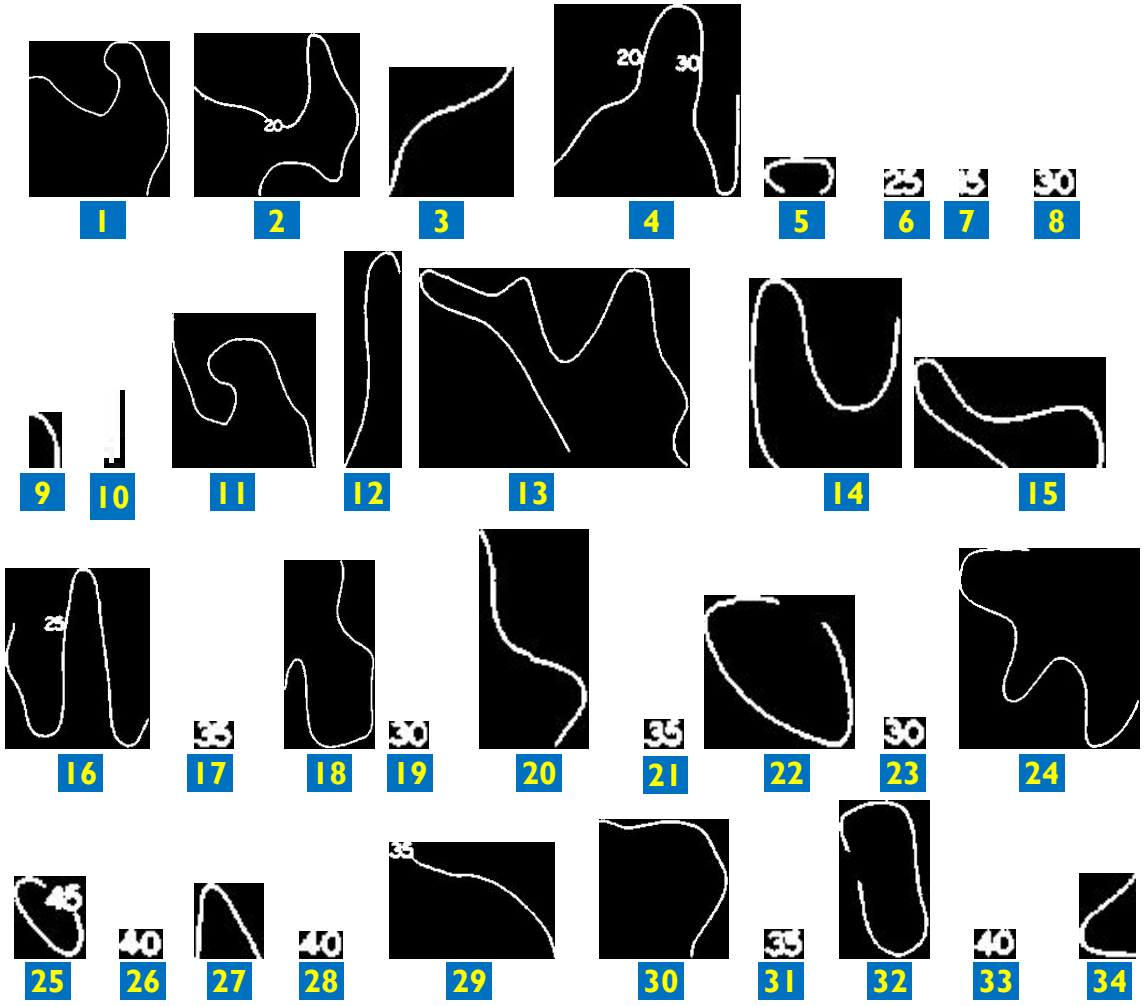


Altitude tags



Contours with attached text

Region properties Extraction



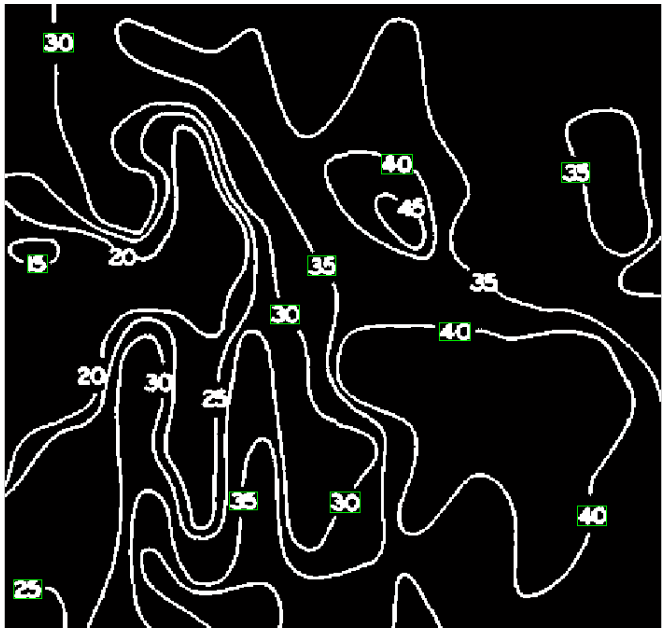
CC	MBR Area	Aspect Ratio	Density
1	33998	1.07	30.52
2	28380	1.04	20.11
3	4290	1.02	19.32
4	22500	1.00	15.77
5	703	1.95	4.29
6	260	1.54	1.51
7	182	1.08	1.43
8	273	1.62	1.55
9	464	1.81	5.21
10	42	4.67	1.05
11	26394	1.04	24.46
12	11385	3.76	18.45
13	45180	1.39	25.95
14	7840	1.23	13.18
15	5757	1.77	11.33
16	20410	1.21	15.27
17	260	1.54	1.53
18	21630	2.04	21.42
19	260	1.54	1.65
20	6384	1.96	17.88
21	280	1.43	1.68
22	6162	1.01	11.69
23	315	1.40	1.83
24	34547	1.08	30.25
25	1517	1.11	3.81
26	308	1.57	1.68
27	1404	1.08	6.47
28	286	1.69	1.71
29	12285	1.48	24.47
30	13334	1.04	23.98
31	280	1.43	1.68
32	6572	1.71	11.10
33	294	1.50	1.79
34	1290	1.43	6.94

Step 2: Expectation Maximization(EM) Algorithm

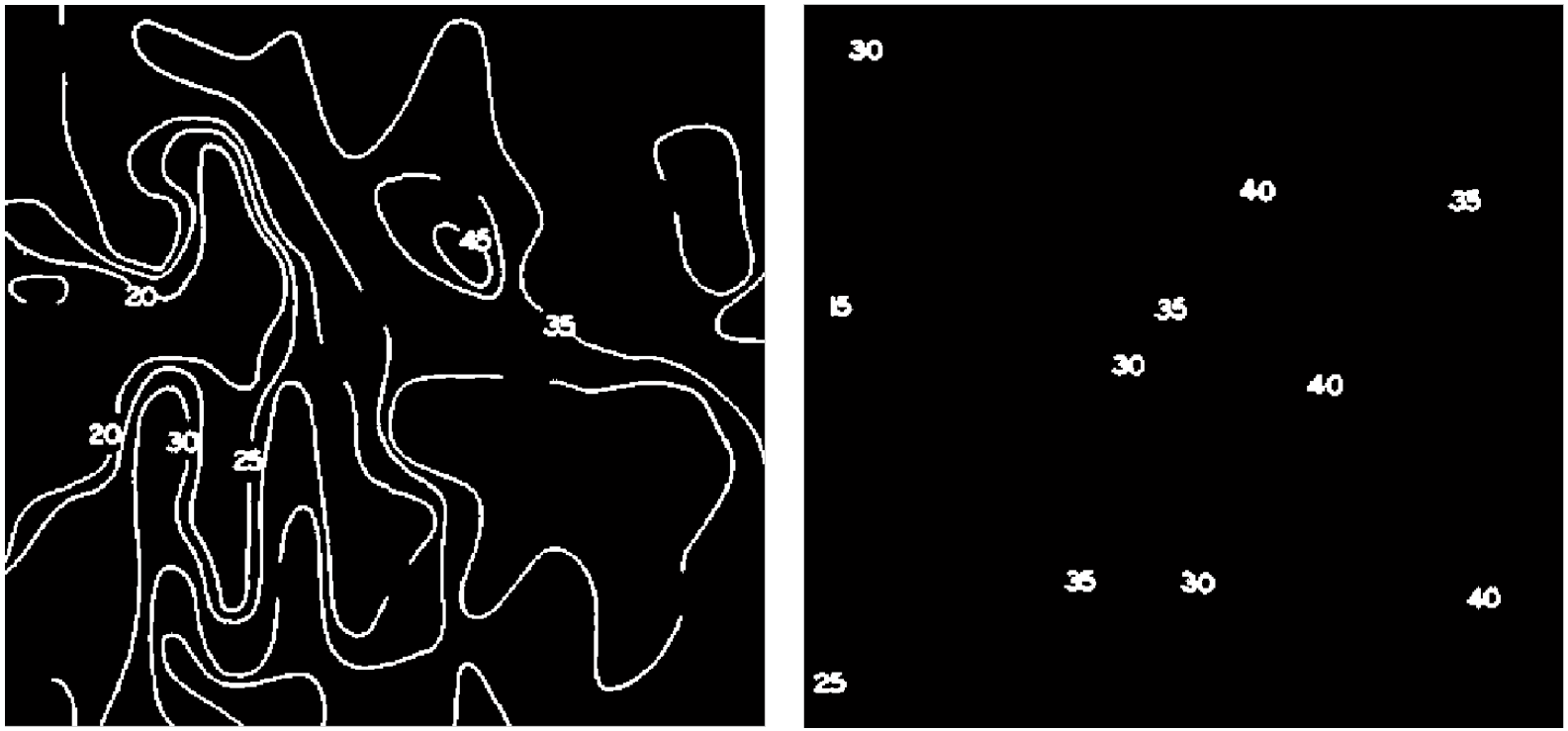
- The EM algorithm is an iterative algorithm for calculating the maximum-likelihood or maximum-a-posteriori estimates of finite mixture models, when the observations can be viewed as incomplete data i.e. when there is a many to one mapping from underlying distribution to the distribution governing the observation.
- Each iteration of the algorithm consists of an expectation step(E-Step) followed by a maximization step(M-Step). The E-Step is with respect to the unknown underlying variables, using the current estimate of parameters and conditioned upon the observation, the M-Step then provides new ML estimates/updates of the parameters given in the E-Step. These two steps are iterated until convergence.
- To use EM Algorithm, the number of component densities 'C' in the mixture has to be known and an initial guess for the value of component parameters has to be made. Once initial estimates are found, the parameter estimates are updated using iterative EM update equations.

Step 3: Filter text not connected to contours

- Pixel Density (Area of MBR/Total Object pixels) and Aspect ratio are used.
- EM is used to cluster altitude tags not connected to contours and contours
- Data is divided into two (C) equal groups, and $\mu_k^{(0)}, \sigma_k^{(0)}$ initialized to mean and standard deviation of each of these groups respectively.
- After EM is run, each data point is assigned a posterior probability, based on which each component is classified as belonging to either the text or contour.



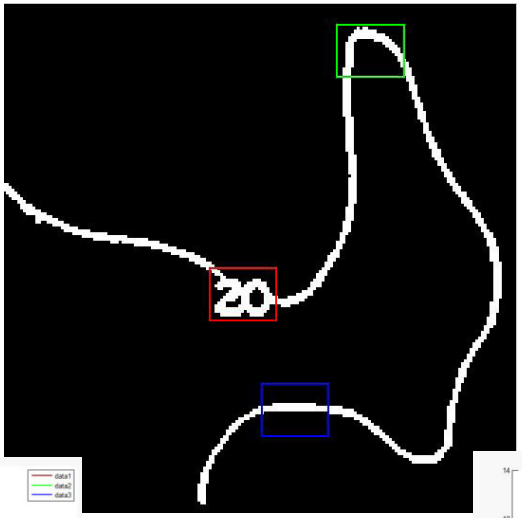
Filter tags not connected to contours

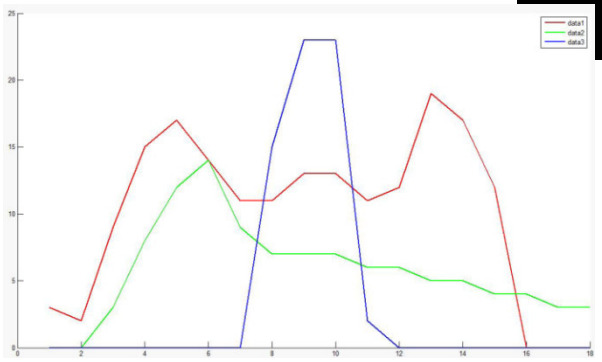


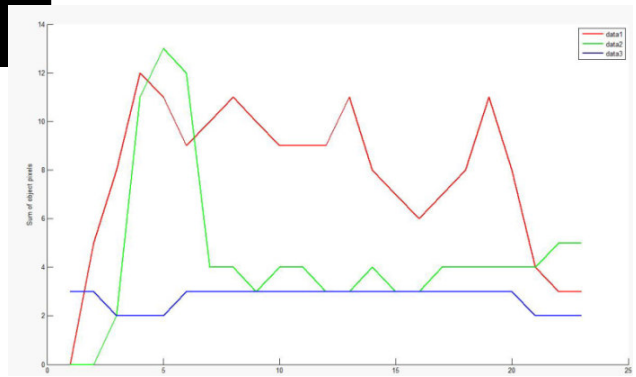
Output of Stage I

Step 4: Text connected to contours

- As region properties fail to classify, a more local approach taking into account the spatial characteristics of the image is essential.
- Different binary regions of the segment have been enclosed by three rectangles of same size, area enclosed in red is the text, area in green is the contour at a bend (two segments occur in the same window) and blue is a single contour segment.
- The row and column projections of the areas enclosed are shown in figure.



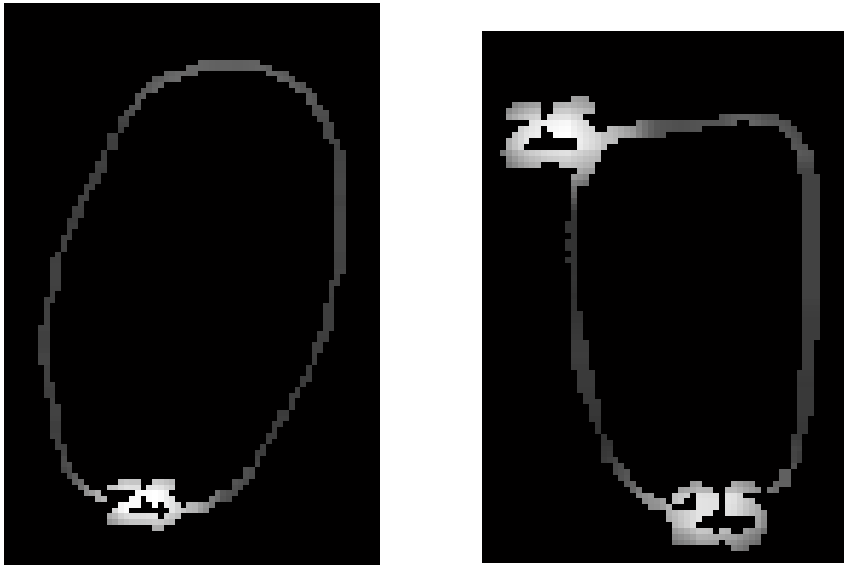




Row & Column projections of regions of text and contours in a component

Step 4: Text connected to contours

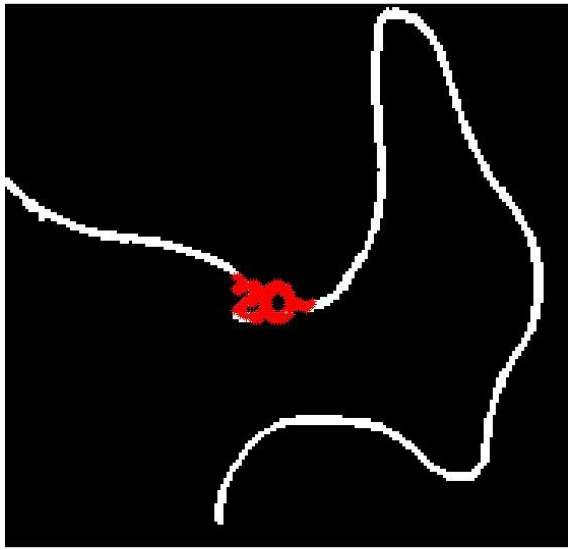
- The maximum of all the filtered MBRs dimensions is stored in w, and average pixel count is stored in T_{pc}.
- The following procedure is followed on each component which does not belong to the text.
  - The locations of all the object pixels in the component are stored and at each object pixel, a neighborhood of size w is considered.
  - Total count of the pixels in the window, i.e. area under the Projection curve, is computed.
  - Each pixel is replaced by the total count of pixels around a window, centred at that pixel.



Projection Images of contours with text attached

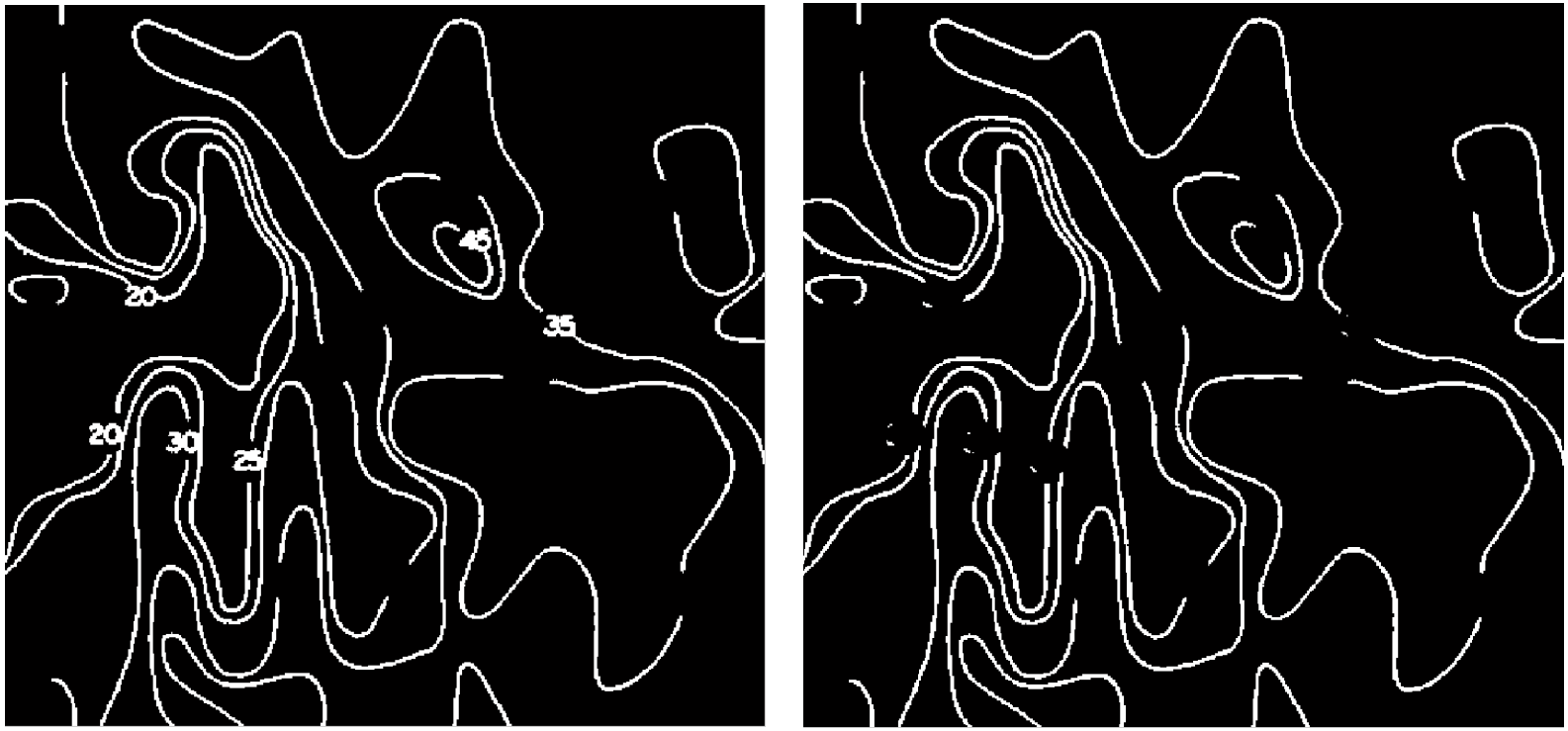
Filter text connected to contours

- Pixels where text is present are brighter, and the histogram has values beyond a threshold, when compared to contour without text.
- If the count is greater than threshold, T_{pc}, it belongs to the text and hence is made equal to background.(zero in this case)

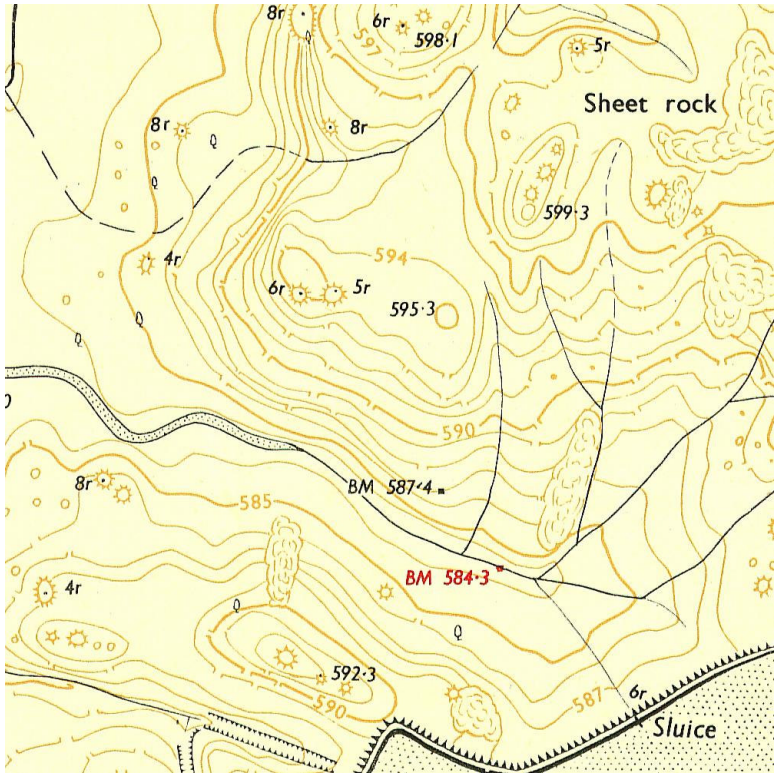


Pixels with higher intensities

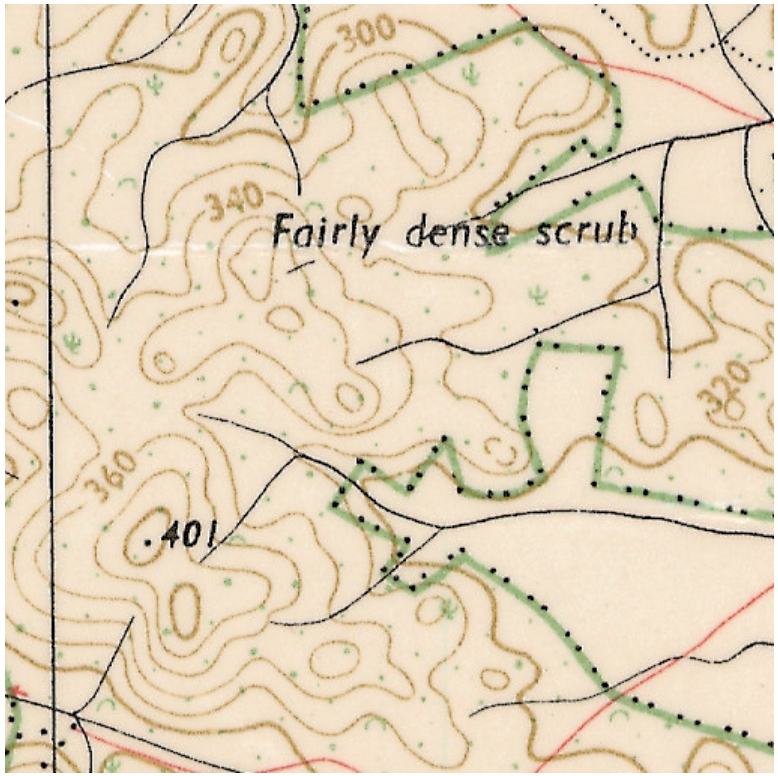
Output after Stage II



Sample Maps

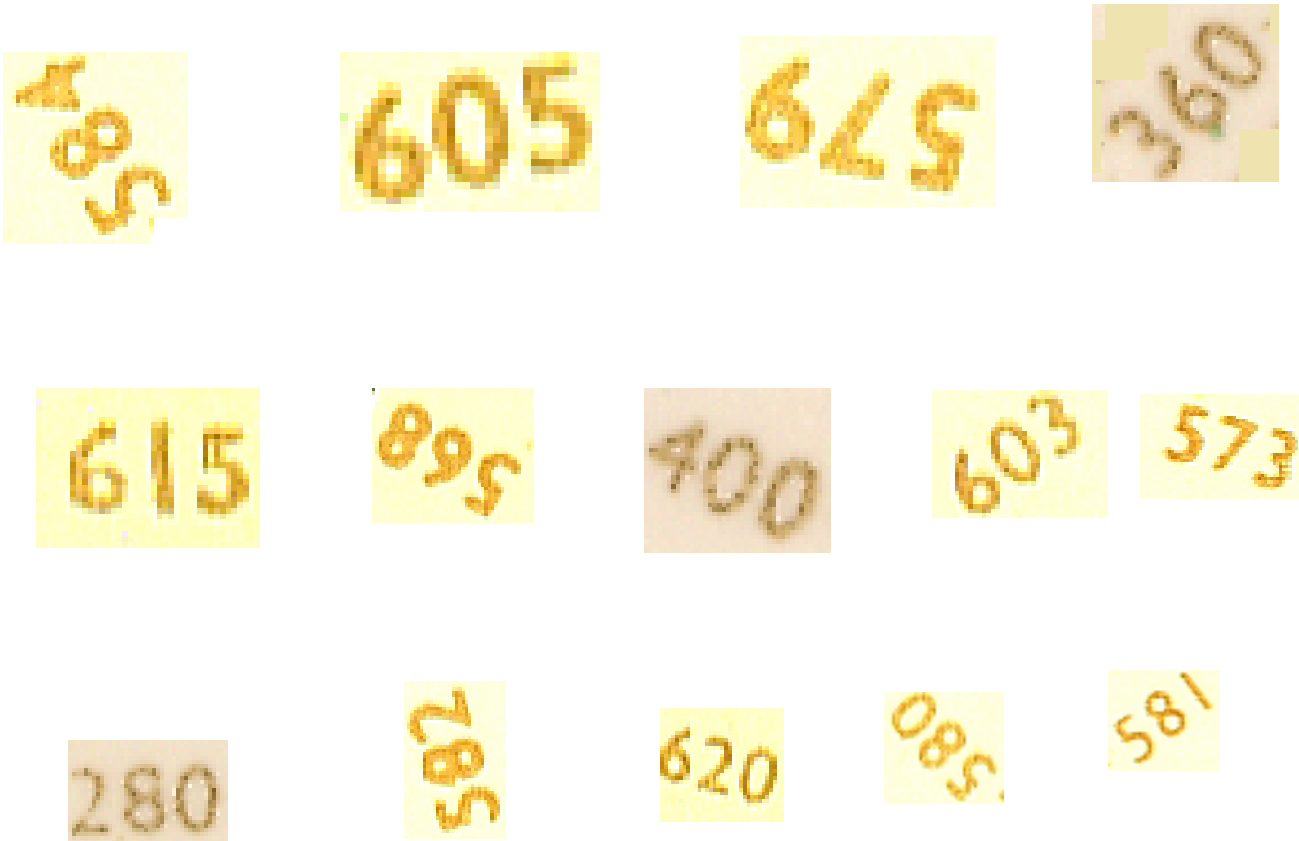


UOH Map



Scanned Map

Sample Images (cropped from previous slide)



Issues

- Altitude tags do not belong to any standard font, the tags are not printed but hand written using markers.
- Altitude tags are not aligned along any axis, but rotated along the contour line to which the value has to be tagged.
- Altitude tags are noisy. As text is extracted from a scanned colour topographic map, gaps are formed due to variations in colour introduced by the scanning process. Also when text attached to contour lines is filtered from the contour layer, some noise gets attached to the tag.

Handwritten digit recognition - Survey

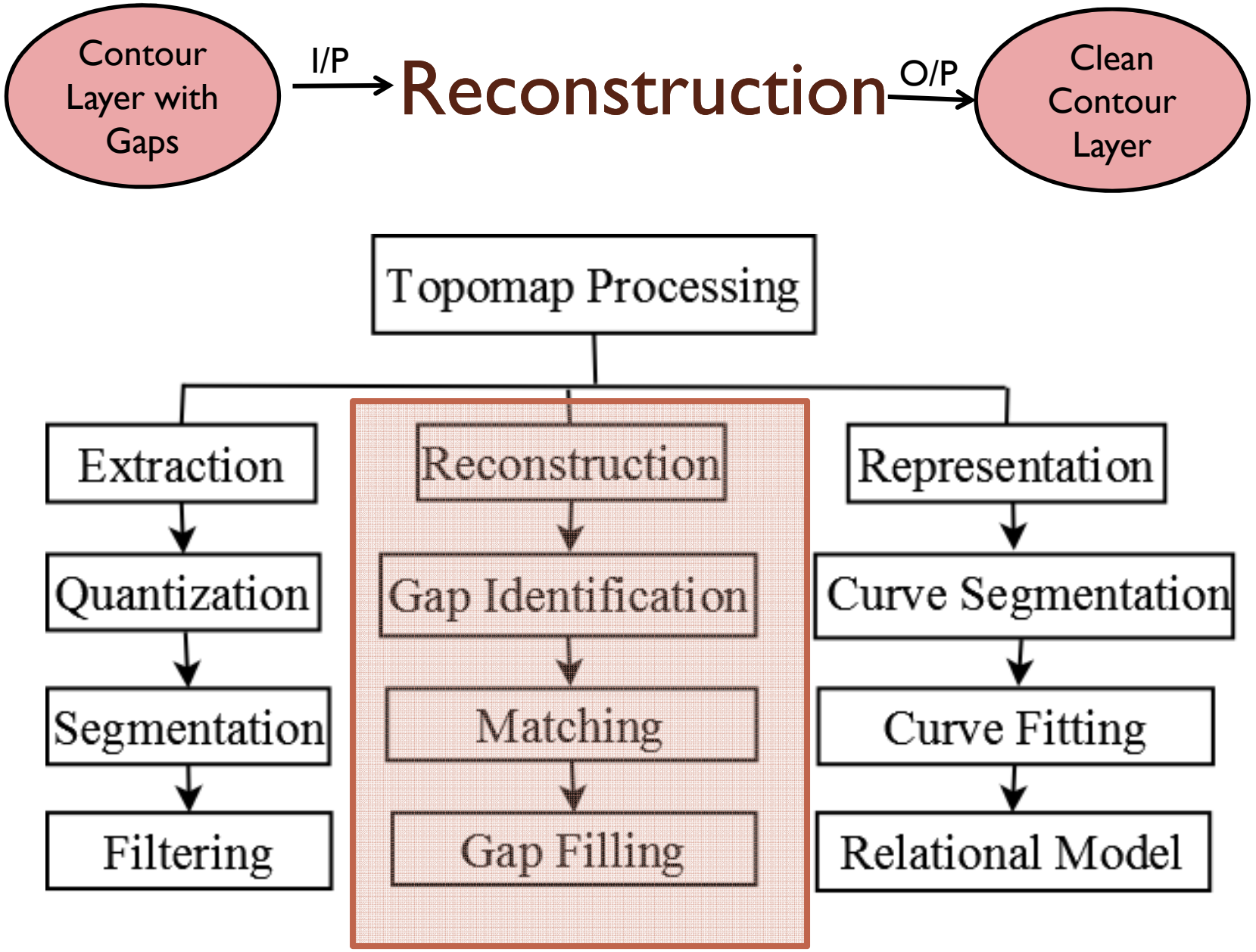
- Handwritten digit recognition is an active topic in OCR applications and pattern classification/learning research.
- Features extracted range from geometric moments to contours and curvatures
- Classification techniques range from template matching to neural networks.
- Databases like CENPARMI , CEDAR, and MNIST have been widely used in classifier design and evaluated in character recognition and classification/learning research
- Different classifiers proved on MNIST database by LeCun had shown recognition rate from 88% till 99.3%

Recognition rate of different classifiers on MNIST database

METHODS	ERROR RATE
Linear classifier	12.0
Linear classifier (nearest neighbor-NN)	8.4
Pairwise linear classifier	7.6
K-NN, Euclidean	5.0
2-layer NN, 300 hidden units (HU) (28x28-300-10)	4.7
2-layer NN, 1000 HU (28x28-1000-10)	4.5
2-layer NN, 1000 HU, [distortions] (28x28-1000-10)	3.8
2-layer NN, 300 HU, [distortions] (28x28-300-10)	3.6
1000 RBF (Radial Basis Function) + linear classifier	3.6
40 PCA (Principal Component Analysis) + quadratic classifier	3.3
3-layer NN, 300+100 HU (28x28-300-100-10)	3.05
3-layer NN, 500+150 HU (28x28-500-150-10)	2.95
3-layer NN, 300+100 HU, [distortions] (28x28-300-100-10)	2.5
3-layer NN, 500+150 HU, [distortions] (28x28-500-150-10)	2.45
K-NN Euclidean, deslant	2.4
LeNet-1 [16x16]	1.7
2-layer NN, 300 HU, [deslant] (20x20-300-10)	1.6
K-NN, Tangent Distance, [16x16]	1.1
SVM (Support Vector Machine) poly 4	1.1
LeNet-4	1.1
LeNet-4 / K-NN	1.1
LeNet-4 / Local	1.1
Reduced Set SVM poly 5	1.0
LeNet-5	0.95
Virtual SVM poly 9 [distortions]	0.8
LeNet-5 [distortions]	0.8
Boosted LeNet-4 [distortions]	0.7

Recognition of digits

- The convolutional neural network created by LeCun et.al has been used for the recognition of digits in altitude tags.
- This network provides some degree of shift, scale and distortion invariance as their basic architectural ideas include local receptive fields, shared weights (or weight replication), and spatial or temporal sub sampling.
- The convolutional neural network is trained using MNIST database.


Contour Layer with Gaps
I/P
Reconstruction
O/P
Clean Contour Layer
Topomap Processing
Extraction
Quantization
Segmentation
Filtering
Reconstruction
Gap Identification
Matching
Gap Filling
Representation
Curve Segmentation
Curve Fitting
Relational Model

Reconstruction

- Contour line layer is obtained from clustering and some post processing like filtering.
- Gaps are caused when intersecting/overlapping objects on contour lines are filtered.

Steps:

- Find the location of gap and the end points of contour segments contributing to gap.
- Pair/Match the contour segments of each gap.
- Fill the gaps in the contour lines.

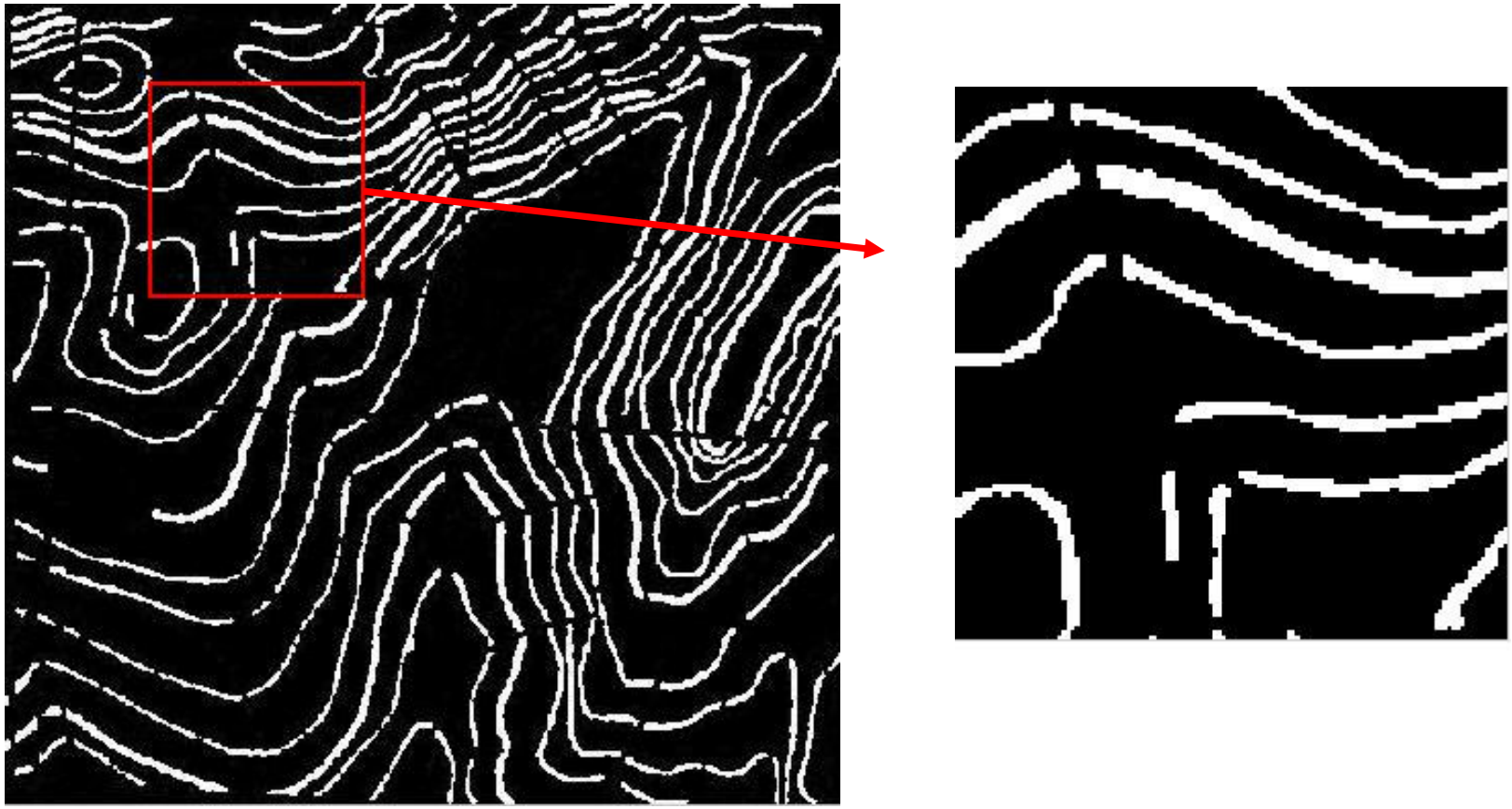
Gap Identification

- The algorithm for Gap Identification operates by reducing the binary contour image with gaps to one consisting of segments of those contours which contribute to gaps.
- An additional feature of our algorithm is finding the end points of the segments. The end points play an important role in matching of contour segments and gap filling.

Gap Identification

- Input is an image of size MXM, where M is a power of 2.
- Move a NXN window along the image overlapping N/2 rows or columns.
- In each window, apply connected components and consider each component for further analysis.
- If a curve segment has both the end points on the edges of the window, it cannot contribute to a gap.
- So the curve whose end point does not touch the edge, leads to a gap.
- After finding such segments, an adjacency matrix is prepared by marking 1 between components which have some pixels in common i.e., 'and' ing which does not result in a zero matrix.
- Backtracking from graph theory is used on the adjacency matrix to collapse the segments.

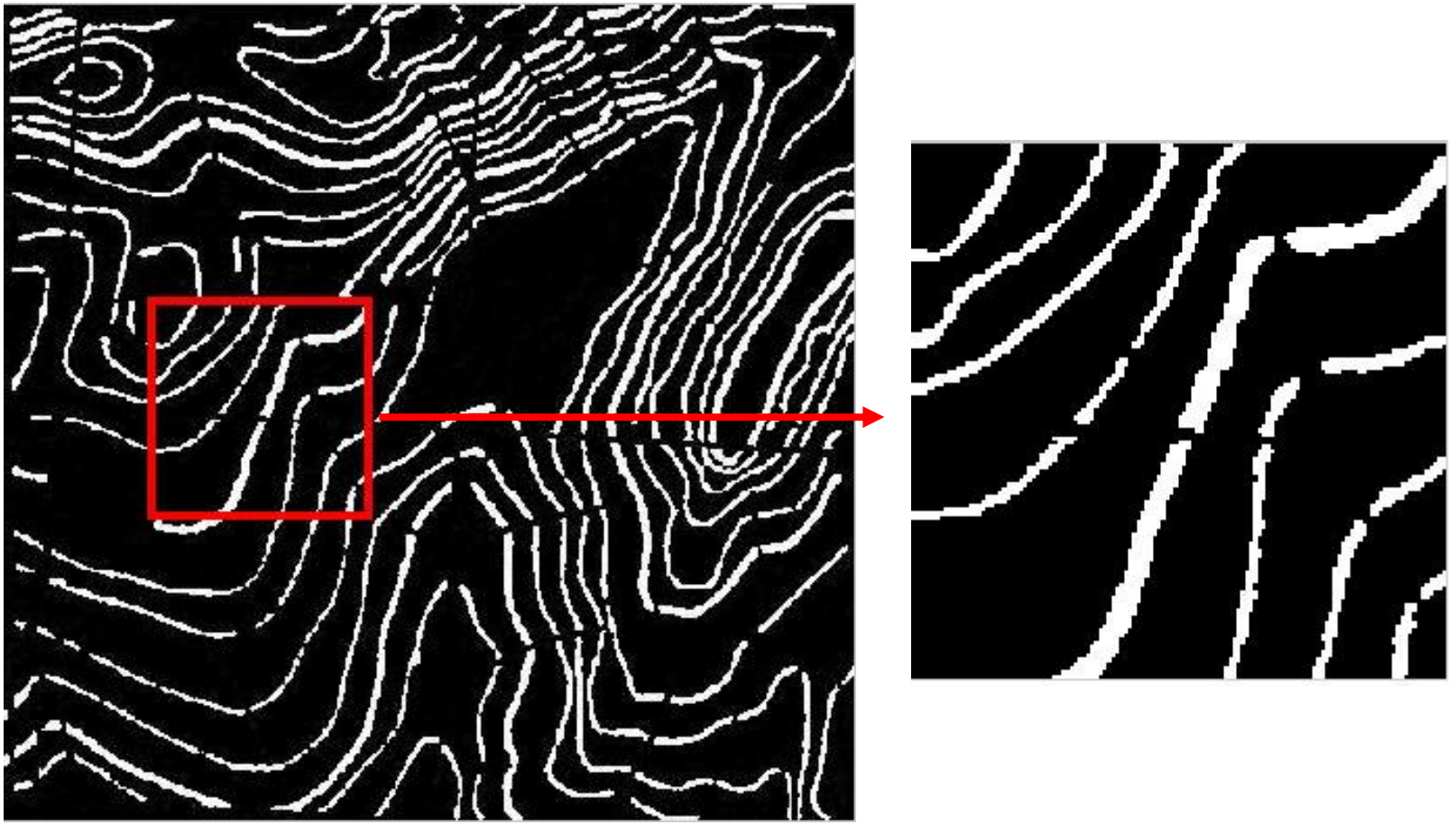
Gap Identification



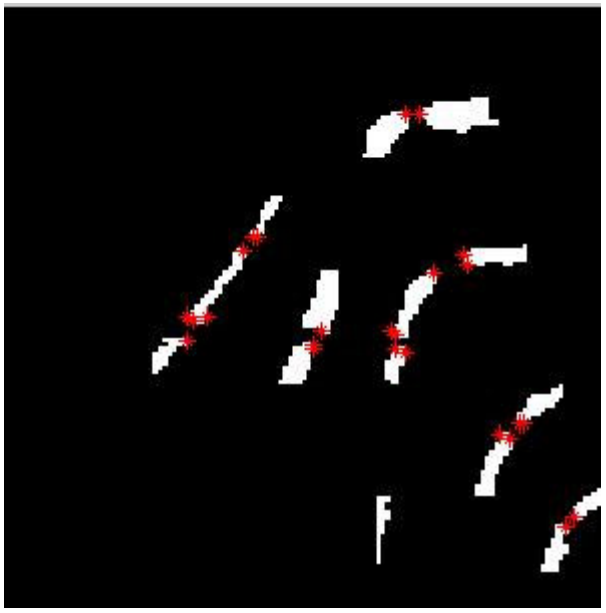
Gap Identification

Segments with gaps and end points

Gap Identification



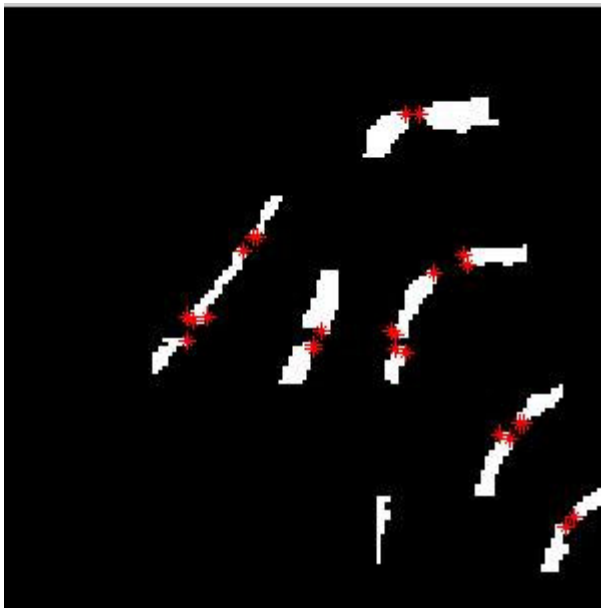
Gap Identification



Pairing Gap Segments

- Once the contour segments in the vicinity of a gap have been isolated, pairs of contour segments which are supposed to close the gap are found by matching.
- To reduce the complexity of pairing it is required first to identify potential end points of the segments contributing to a given gap. This is possible by classifying pixels in the gap and surrounded by the gap either to the contour or background i.e. the considered pixels should be clustered to two clusters, one belonging to the contour and other to the background.
- This clustering problem (2 clusters) has been attempted by EM algorithm.
- Rules are applied to pair segments in the reduced set of contours

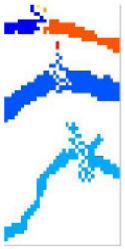
Localization



Local Pairing

- From the extraction stage, when contour layer is filtered from topomap, pixels of all other layers are grouped to obtain 'other layer'
- some object pixels in the other layer should belong to the gap. An image is obtained by 'AND'ing cloud pixels and 'other image' , which has pixels of cloud which also belong to the other layers(say A_l)
- when A_l is 'OR'ed with contour image, an image which fills few gap pixels thereby connecting some contours is obtained

(a) Component of clouded GC Image (b) Contour Image (c) Other layer Image (d) Pixels of cloud and other layer (e) Result Image

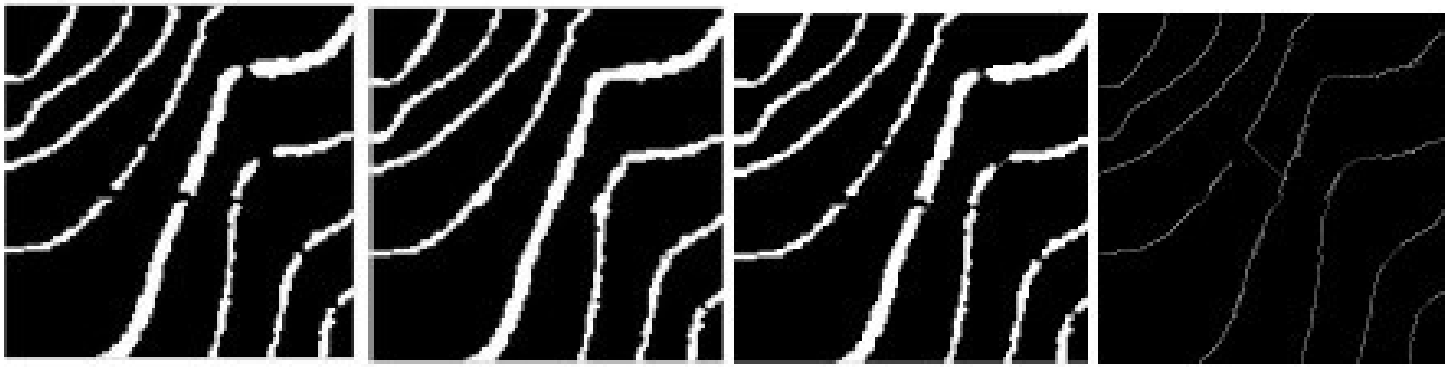


(f) CC of Result image

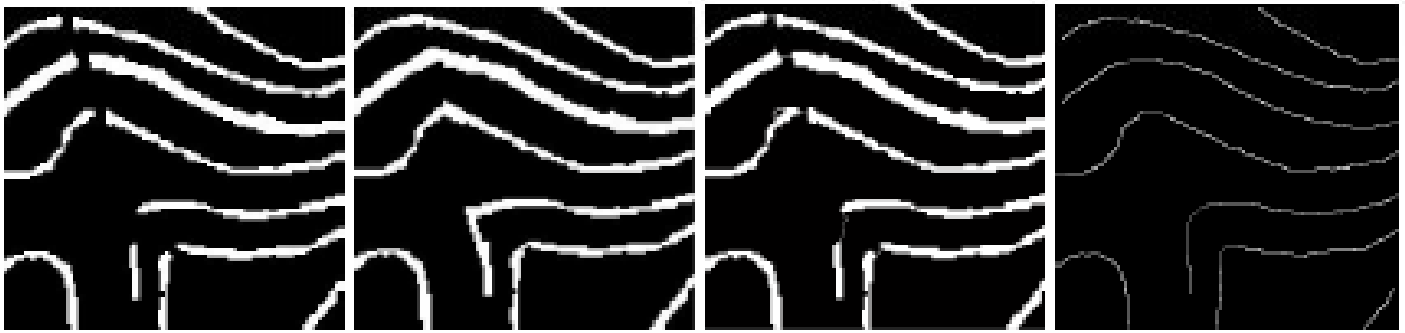
Gap Filling

- Gap filling is the process of identifying the background pixels between two contour segments which should belong to the object(contour) so that the contour line is either closed or ends at the boundary of topomap.
- Gap filling is attempted using two different approaches:
  - Weaving : This approach uses image properties and takes into consideration the thickness and orientation of the contour segments.
  - Bezier approximation: The contour segments have been approximated using Bezier curves and hence the gap between two such approximated contour segments is filled using the Bezier coordinates.

Results

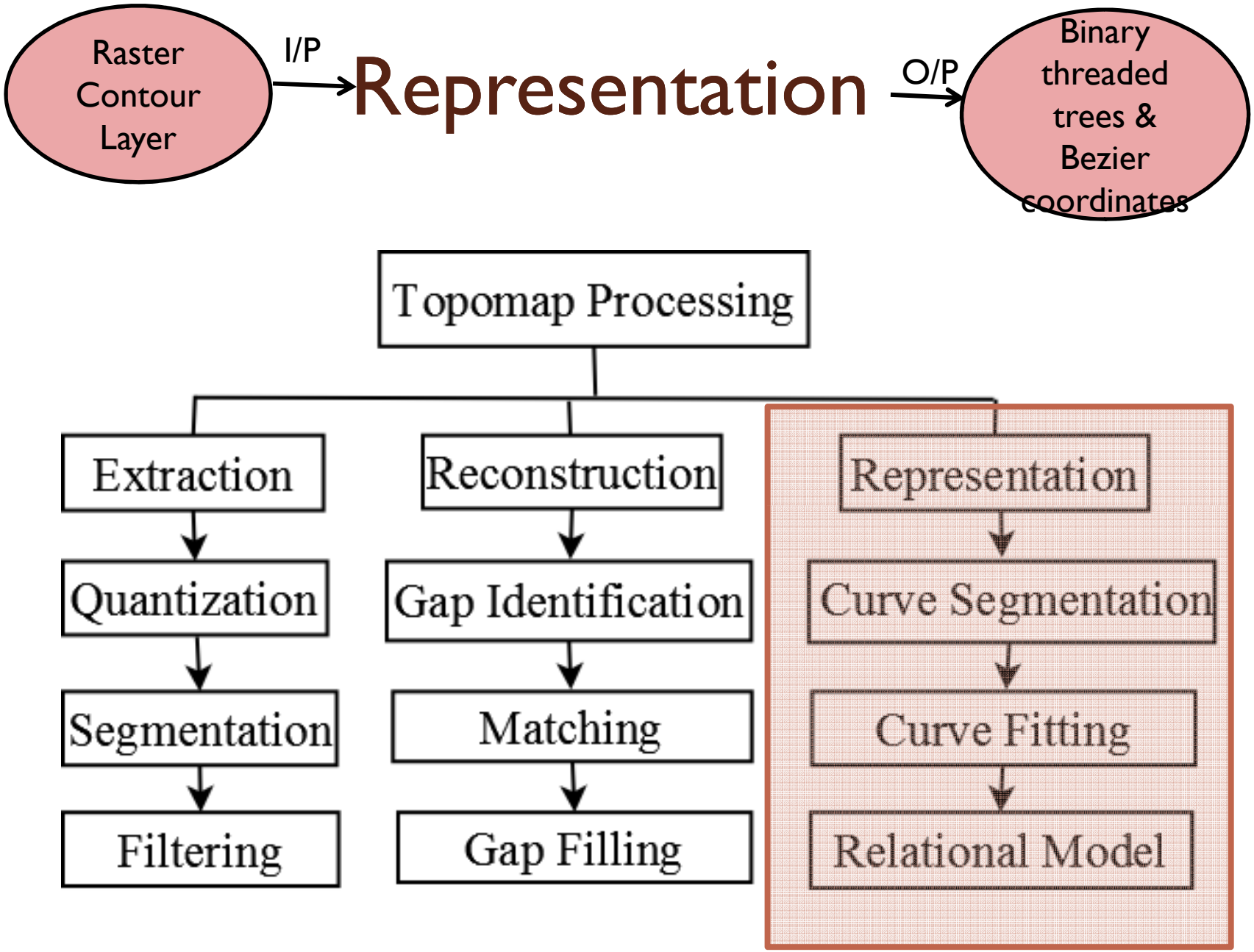


(a) Sample Image (b) Weaving result (c) Bezier interpolation result (d) Global reconstruction result((JS07)



(e) Sample Image (f) Weaving result (g) Bezier interpolation result (h) Global reconstruction result((JS07)

- Bezier interpolation and Global reconstruction algorithm, only fill the gap with single pixel width line whereas Weaving fills the area of gap with a thick line.
- Global reconstruction requires thinning prior to applying the gap filling.
- Gap filling using Bezier interpolation can be an on-line process.


Raster Contour Layer
I/P
Representation
O/P
Binary threaded trees & Bezier coordinates
Topomap Processing
Extraction
Quantization
Segmentation
Filtering
Reconstruction
Gap Identification
Matching
Gap Filling
Representation
Curve Segmentation
Curve Fitting
Relational Model

Contour layer Representation

- Storing raw images of contour layer cannot be used for subsequent processing steps like vectorization, 3-D modelling.
- Since these curves are arbitrary in shape, they cannot be described by finite mathematical expressions.
- Hence an efficient, stable and robust representation scheme has to be formulated.

A good representation should have the following features:

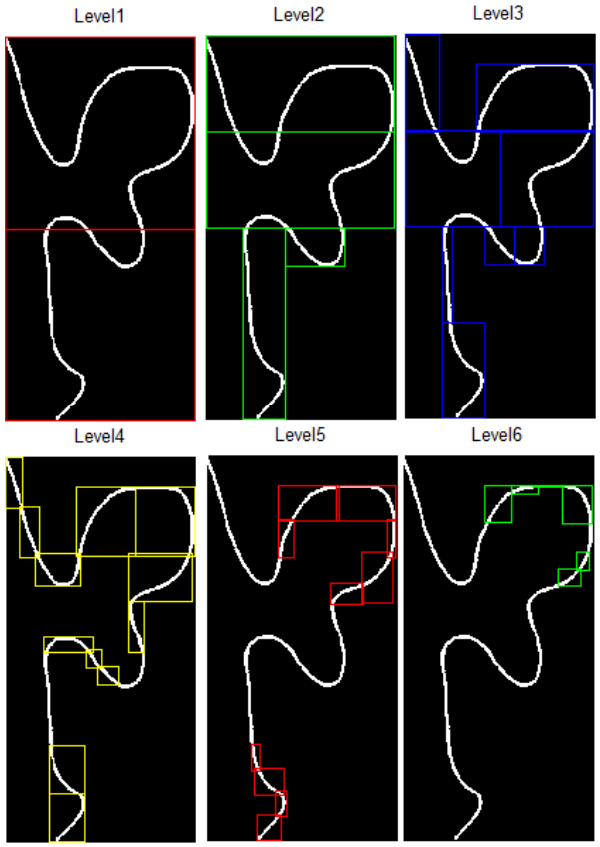
- It has to handle various complex operations like querying, set operations, change detection etc.
- It should support visualization which involves various transformations like scaling, rotation. Hence it should support multiple resolutions and should be invariant to affine transformations.
- Topomap images, owing to their large size, require a parallel computational architecture for processing. Therefore, a good representation of contour layer should be amenable to parallel implementation.
- The representation has to be effective in addressing both region level and pixel level operations to be performed on the contour layer.

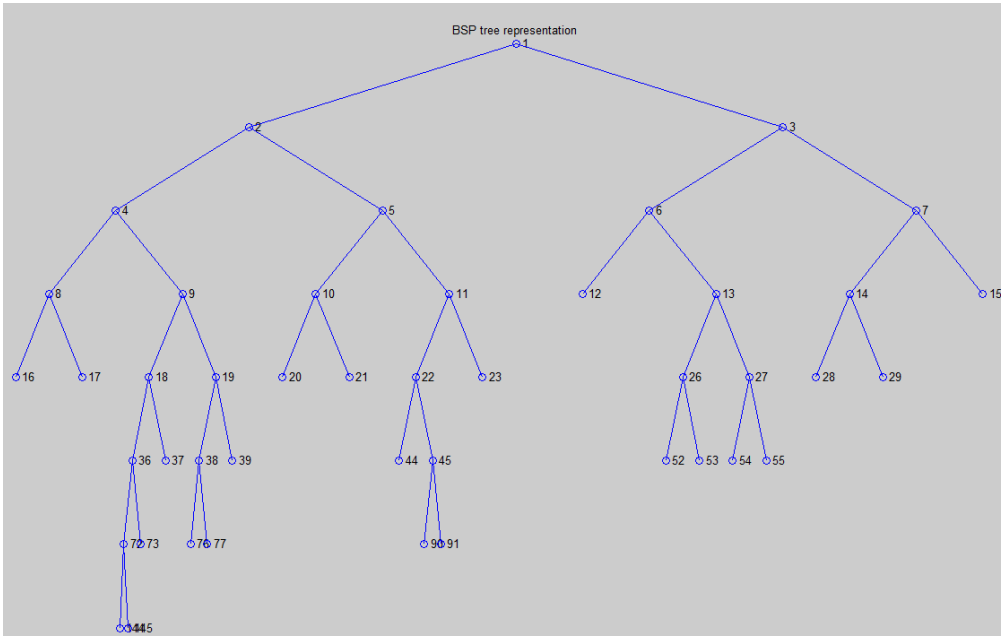
Curve Segmentation

Connected component analysis is applied on the binary contour layer obtained from Extraction stage of topomap processing. Each component of CCA gives one contour segment/curve of the contour layer. The following procedure is applied on each contour segment/component of the contour layer:

1. The curve is enclosed in a minimum bounding rectangle(MBR).
2. The ratio of number of object pixels in a curve to the area of the enclosing Minimum Bounding Rectangle (MBR), density ratio, is used as the measure of complexity of the shape of the curve.
3. A curve is split if the ratio is less than a specified threshold.
4. The direction of split is along the minimum dimension of MBR. In other words, the curve is split vertically if the length of MBR is more than or equal to the height and horizontally if the height is greater than length.
5. Steps 1 - 4 are repeated on each segment of the curve, until the density ratio for all the segments is more than the threshold value.
6. Steps 1 - 5 are repeated for each component of the contour layer.

Results

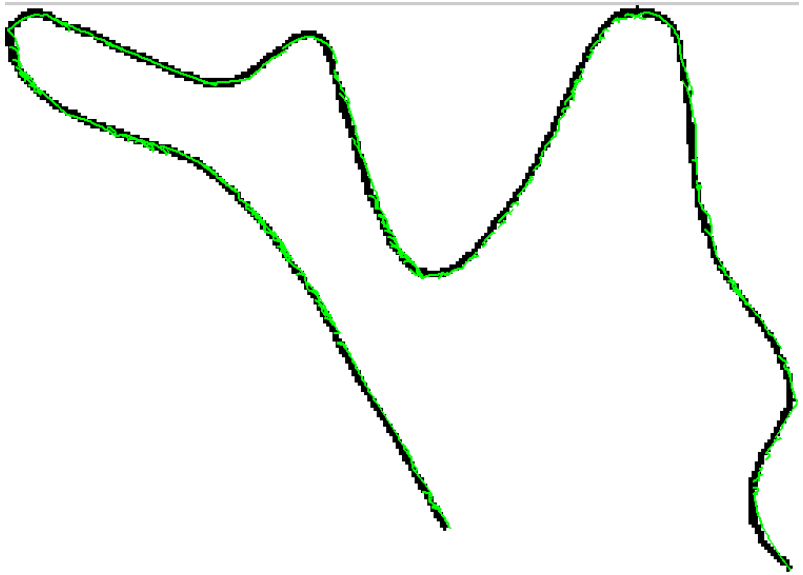

Level1
Level2
Level3
Level4
Level5
Level6


BSP tree representation
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
26
27
28
29
36
37
38
39
44
45
52
53
54
55
72
73
76
77
90
91

Curve Approximation

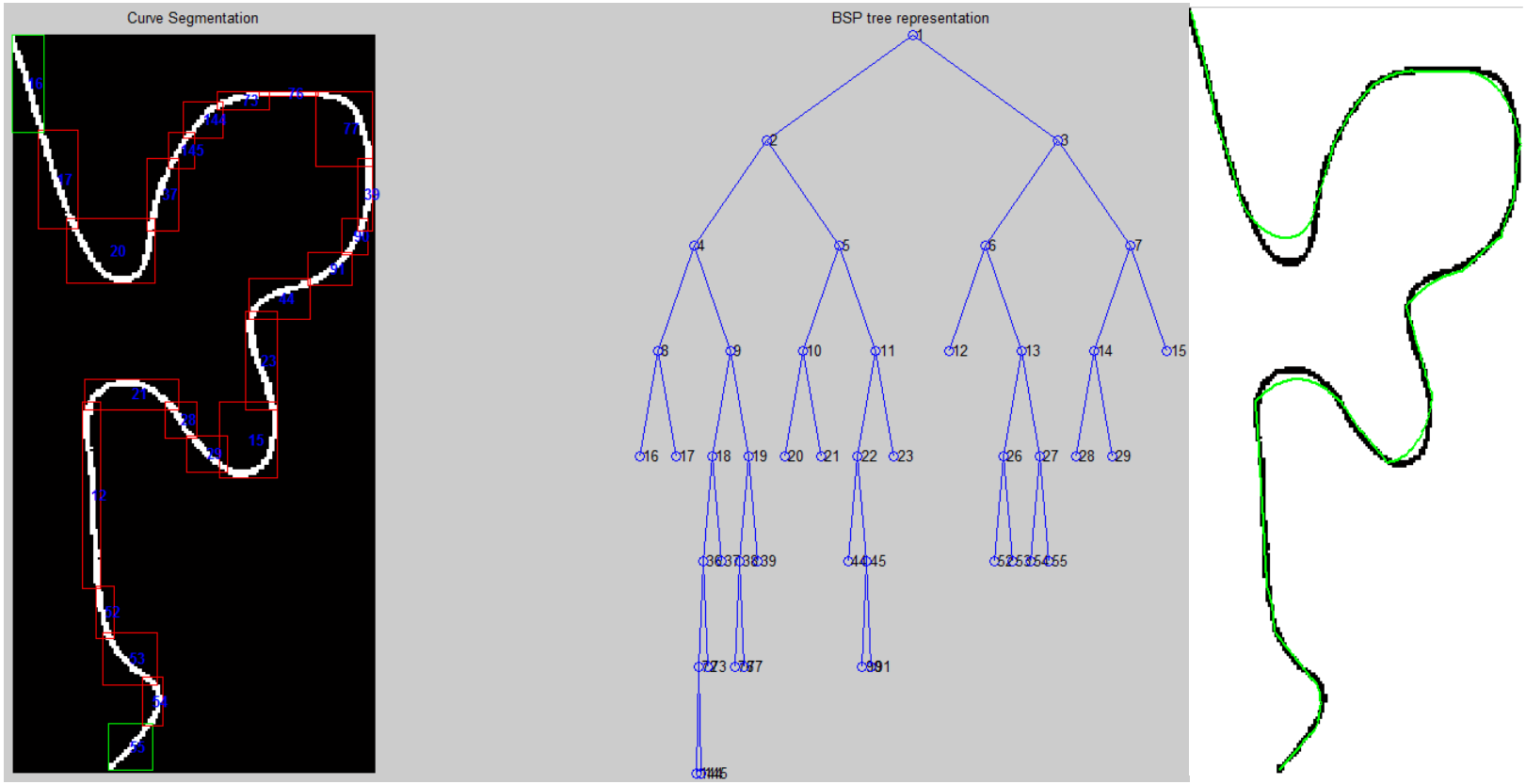
- Each region of the segmented contour contains portions of curve, with simple shapes and sizes compared to the original curve. Boundary-following can therefore be employed on the segment in each of these regions, and hence can be represented using various contour based approaches.
- Boundaries can be represented efficiently with various spline models like Bezier splines, B-splines, Hermite interpolation and rational cubic interpolation.
- In addition to high data reduction this approach has various other advantages like translation, rotation, scaling, and deformation of shapes without losing any quality
- Bezier curves are piecewise polynomial functions that can provide local approximation of contours using small number of parameters.
- A cubic Bezier has been chosen as it provides a good trade-off between complexity and accuracy.

Results



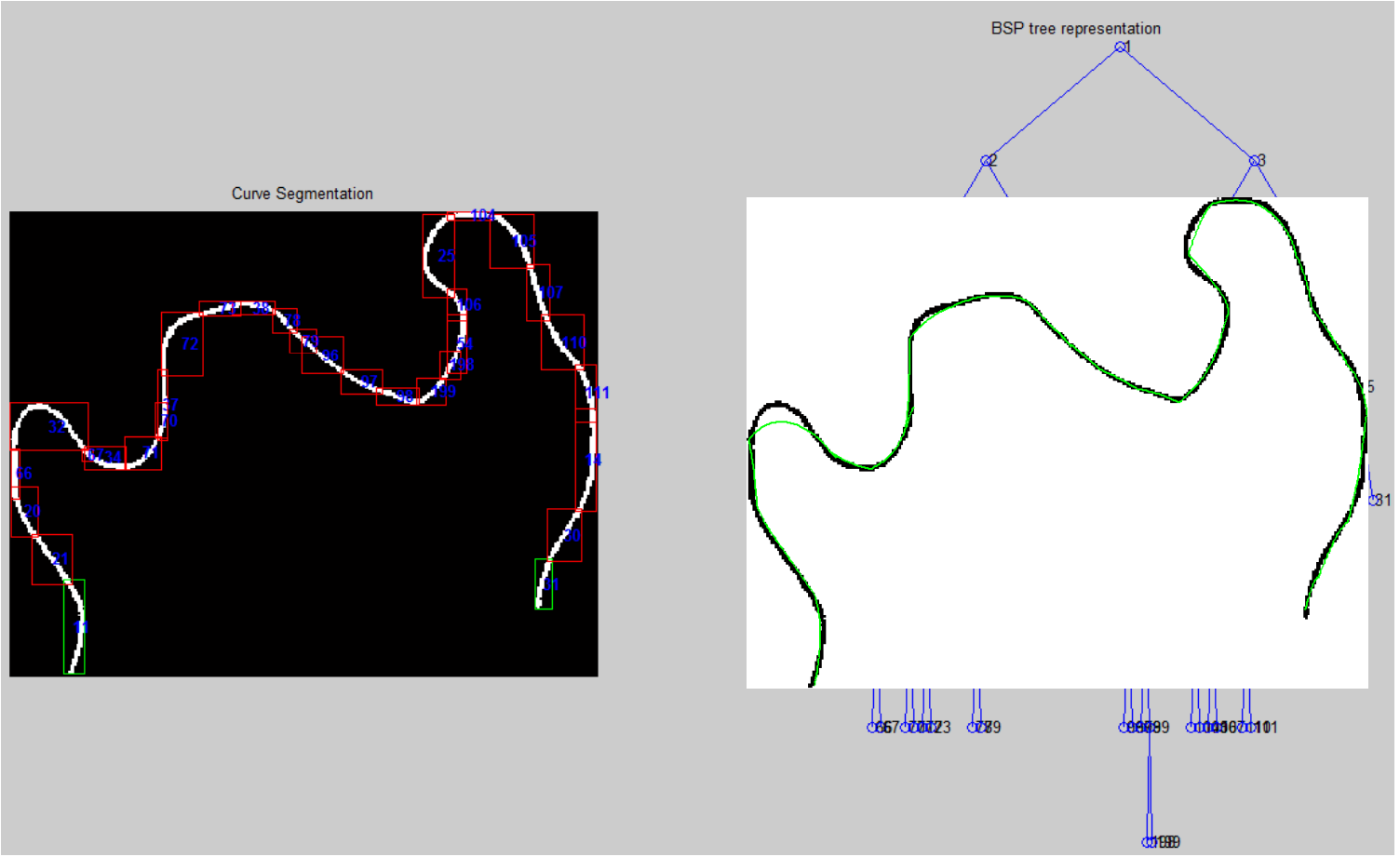
P0x	P0y	P1x	P1y	P2x	P2y	P3x	P3y
151.5	237.5	137	246	150.5	239.5	165.5	239.5
151.5	237.5	130	249	120.5	249.5	115.5	248.5
91.5	232.5	98.5	240	106.5	244.5	115.5	248.5
71.5	222.5	78.5	226	85.5	230.5	91.5	232.5
49.5	218.5	55.5	220	63.5	221.5	71.5	222.5
24.5	215.5	33.5	218	41.5	219.5	49.5	218.5
24.5	180.5	4.5	194	6.5	212.5	24.5	215.5
24.5	180.5	24.5	183	26.5	180.5	28.5	180.5
38.5	174.5	32.5	178	26.5	180.5	24.5	180.5
54.5	166.5	48.5	170	43.5	172.5	38.5	174.5
68.5	156.5	62.5	162	57.5	165.5	54.5	166.5
80.5	147.5	75.5	152	69.5	155.5	68.5	156.5
78.5	126.5	84.5	135	83.5	142.5	80.5	147.5
78.5	126.5	77.5	126	80.5	128.5	84.5	131.5
42.5	110.5	53.5	115	66.5	120.5	78.5	126.5
28.5	106.5	32.5	110	39.5	111.5	42.5	110.5
14.5	88.5	10.5	98.5	17.5	104.5	28.5	106.5
25.5	63.5	24.5	72.5	20.5	80.5	14.5	88.5
12.5	32.5	16.5	43.5	21.5	55.5	25.5	63.5
8.5	1.5	8.5	26.5	7.5	18.5	12.5	32.5
12.5	32.5	43.5	44.5	46.5	55.5	54.5	65.5
54.5	65.5	55.5	69.5	60.5	75.5	65.5	78.5
65.5	78.5	68.5	82.5	73.5	86.5	79.5	89.5
79.5	89.5	79.5	90.5	86.5	95.5	93.5	100.5
93.5	100.5	99.5	104	107.5	107.5	114.5	110.5
114.5	110.5	117	113	122.5	115.5	128.5	119.5
128.5	119.5	135	124	141.5	127.5	148.5	131.5
148.5	131.5	149	132	158.5	136.5	167.5	140.5

Results



Leaf Nodes are connected as:
55->54->53->52->12->21->28->29->15->23->44->91->90->39->77->76->73->144->145->37->20->17->16

Results



Spatial query analysis

- Point Query: Given a query point P and a set of objects M. The point query yields all the objects of M geometrically containing P.

  Example Q1 : The nearest contour line to a point. This is useful in obtaining the altitude value at that point.

- Region Query: Given a polygonal query region R and a set of objects M, the region query yields all the objects of M sharing points with R.

  Example Q2: What is the range of altitude values of selected region or the range of altitudes through which a river flows.

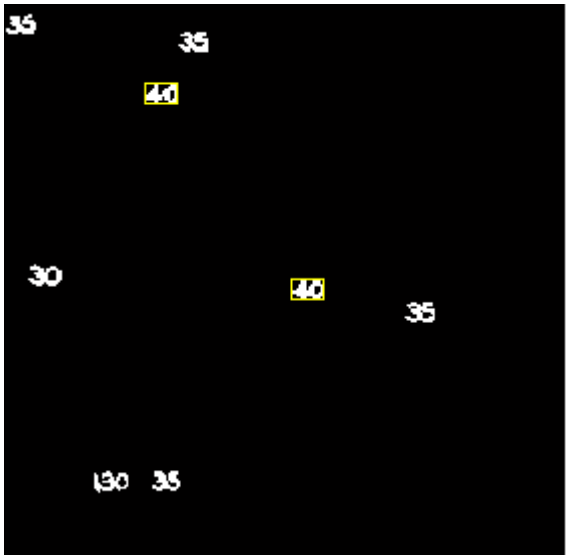
- Spatial Join: For two given object sets A and B the spatial join operation yields all pairs of objects whose spatial components intersect.

  Example Q3: Given a contour, return the various regions through which it passes.

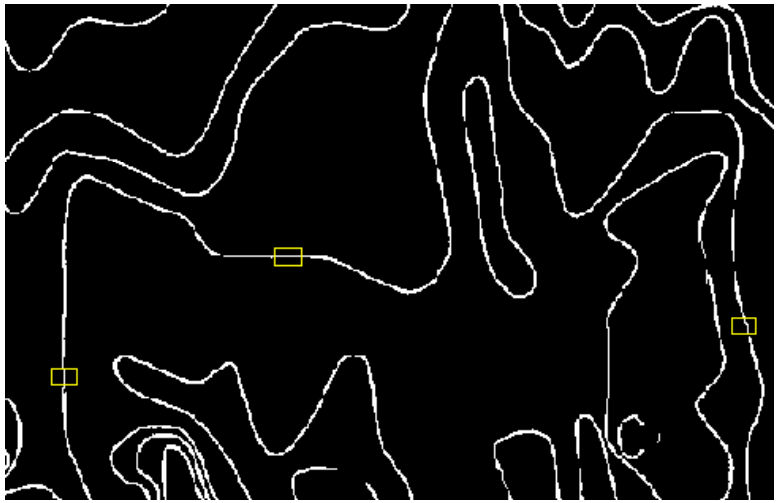
Tagging Altitude Values

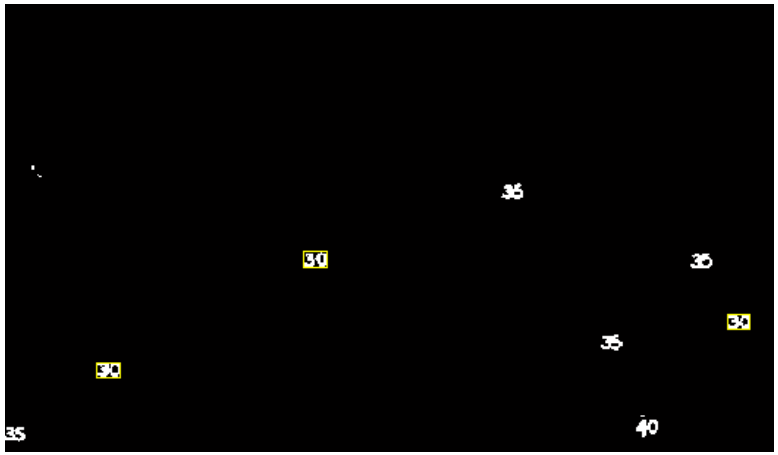
- The contour layer has been reconstructed and represented using a binary threaded tree and each segment approximated using Bezier curves.
- The altitude values filtered from the contour lines have been recognized using LeNet - 5.
- To generate a 3D model from the contour lines, the altitude values have to be tagged to the contour lines.
- The altitude values can be tagged to the contour lines using tree representation.
- The MBR coordinates of labels are known from the text layer.
- The contours which intersect with these labels can be found by searching trees, which store MBR coordinates of the segments of contours which is similar to Region query, explained earlier .
- When more than two contours are obtained, the contour whose tree intersects with the MBR of altitude tag at a deeper level is chosen.

Results



Results





Publications

- Abstract on "Topographic map processing on GEON", Proceedings of AP Science Congress, 2008
- "Automatic gap identification towards efficient contour line reconstruction in topographic maps." Asia International Conference on Modelling and Simulation, p:309 - 314, 2009
- "Adaptive Altitude Feature Recognition for Paper based Topographic map", The 4th Mahasarakham International Workshop on Artificial Intelligence, Thailand, 2010
- "Contour Layer Extraction From Colour Topographic Map by Feature Selection Approach", IEEE Symposium on Computers & Informatics, Malaysia, 2011
- "Representation of Contour lines Extracted from Topographic map", communicated to Computers & GeoSciences, Elsevier

References

[AAC09] Jeffrey Armstrong, Maher Ahmed, and Siu-Cheung Chau. A rotation invariant approach to 2d shape representation using the hilbert curve. ICIAR 2009, LNCS, 5627:594 - 603, 2009.

[AK] C.P.Lo Albert and K.W.Yeung. Concepts and techniques of Geographic Information Systems. PH Series in Geographic Information Science, Prentice-Hall, India.

[AS99] Patrice Arrighi and Pierre Soille. From scanned topographic maps to digital elevation models. Proceedings of Geovision99, International Symposium on Imaging Applications in Geology, 1999.

[ASL90] M. Ansoult, P. Soille, and J. Loodts. Mathematical morphology: A tool-for automated gis data acquisition from scanned thematic maps. Photogramm.Eng. Remote Sens, 56(9):1263 - 1271, 1990.

[Bal81] Dana H. Ballard. Strip. trees:A hierarchical representation for curves. Graphics and Image Processing, Communications of the ACM, 24(5):310 - 321,1981.

[BK92] B.K.Ray and K.S.Ray. An algorithm for detection of dominant points and polygonal approximation of digitized curves. Pattern Recognition Letters, 13:849{856, 1992.

[CJS01] H.D. Cheng, *X.H. Jiang, Y. Sun, Jingli Wang,"*Color image segmentation: advances and prospects", Pattern Recognition 34 (2001) 2259 – 2281

[CR88] C.H.Teh and R.T.Chin. On image analysis by the methods of moments. IEEE Trans. Pattern Anal. Mach. Intell PAMI, 10(4):496- 513, 1988.

[CWQ06] Yang Chen, Runsheng Wang, and Jing Qian. Extracting contour lines from common-conditioned topographic maps. IEEE Transactions on Geoscience and Remote Sensing, 44(4):1048 - 1057, April 2006.

[D89] Rhind D. Why gis? ARC News, 11(3), 1989. CA: Environmental Systems Research Institute,Inc.

[D.S93] D.Sarkar. A simple algorithm for detection of signicant vertices for polygonal approximation of chain-coded curves. Pattern Recognition Letters, 14:959 - 964, 1993.

References - contd

[ELA94] N. Ebi, B. Lauterbach, and W. Anheier. An image analysis system for automatic data acquisition from colored scanned maps. Machine Vision Applications, 7(3):148 - 164, 1994.

[FK88] L.A. Fletcher and R. Kasturi. A robust algorithm for text string separation from mixed text/graphics images. IEEE Transactions on pattern Analysis and Machine Intelligence, 10:910 - 918, 1988.

[Gan06] Pankaj Ganguly. Modied arc tree based hierarchical representation of digital curve. Pattern Recognition Letters, 27:529 - 535, 2006.

[GK79] G.M.Hunter and K.Steiglitz. Operations on images using quadtrees. IEEE Transactions on PAMI, 1(2):145 - 153, 1979.

[GW90] Oliver Gunther and EugeneWong. The arc tree :an approximation scheme to represent arbitrary curved shapes. Comput. Vision Graphics Image Process., 51:313 - 337, 1990.

[HGI95] Luo H.Z, Agam G, and Dinstein I. directional mathematical morphology approach for line thinning and extraction of character strings from maps and line drawings. International Conference on Document Analysis and Recognition(ICDAR), 1:257, 1995.

[I.G82] I.Gargantini. An eective way to represent quadtrees. Comm. ACM, 25(12):905 - 910, 1982.

[IJ84] I.M.Anderson and J.C.Bezdek. Curvature and tangential deection of discrete arcs: a theory based on the commutator of scatter matrix pairs and its application to vertex detection in planar shape data. IEEE Trans Pattern Anal. Machine Intell, 6:27 - 40, 1984.

[JE94] J.C.Perez and E.Vidal. Optimal polygonal approximation of digitized curves. Pattern Recognition Letters, 15:743 - 750, 1994.

[JS07] J.Pouderoux and S.Spinello. Global contour lines reconstruction in topographic maps. Ninth International Conference on Document Analysis and Recognition, 2:779 - 783, 2007.

[JY04] Du Jinyang and Zhang Yumei. Automatic extraction of contour lines from scanned topographic map. IEEE, pages 2886 - 2888, 2004.

[KB90] Alireza Khotanzad and Abdelmajid Bouarfa. Image segmentation by a parallel, nonparametric histogram based clustering algorithm. Pattern Recog

nition, 23(9):961 { 973, 1990.

[KGG02] Rangachar Kasturi, Lawrence O Gorman, and Venu Govindaraju. Document image analysis: A primer. Sadhana, 27(1):3 - 22, February 2002.

[KSL+02] Tombre K, Tabbone S, Peissier L, Lamiroy B, and Dosch P. Text /graphics separation revisited. In Document Analysis Systems V, 2423:615 - 620, 2002.

References - contd

[KZ03] Alireza Khotanzad and Edmund Zink. Contour line and geographic feature extraction from usgs color topographical paper maps. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(1):18 - 31, 2003.

[M.S02] M.Salotti. Optimal polygonal approximation of digitized curves using the sum of square deviation criterion. Pattern Recognition, 35:435 - 443, 2002.

[P.C97] P.Cornic. Another look at the dominant point detection of digitized curves Pattern Recognition Letters, 18(4):13 - 25, 1997.

[PGPG07] Joachim Pouderoux, Jean Christophe Gonzato, Aurelien Pereira, and Pascal Guitton. Toponym recognition in scanned color topographic maps. International Conference on Document Analysis and Recognition(ICDAR), pages 531 - 535, September 2007.

[PT08] Aria Pezeshk and Richard L. Tutwiler. Contour line recognition and extraction from scanned colour maps using dual quantization of the intensity image. IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), pages 173 - 176, March 2008.

[PVL+08] P.P.Roy, Eduard Vazquez, Josep Llados, Ramon Baldrich, and Umapada Pal. A system to segment text and symbols from color maps. 7th International Workshop on Graphics Recognition Recent Advances and New Opportunities: LNCS, 5046:245 - 256, 2008.

[P.Y99] P.Y.Yin. Genetic algorithms for polygonal approximation of digital curves. Int. J. Pattern Recognition Articial Intell, 13:1061 - 1082, 1999.

[P.Y03] P.Y.Yin. Ant colony search algorithms for optimal polygonal approximation of plane curves. Pattern Recognition, 36:1783 - 1797, 2003.

[RC01] Cao R and Tan C.L. Text/graphics separation in maps. Proceedings of 4th IAPR International Workshop on Graphics Recognition, 1:44 -48, September 2001.

[RBM07] Romain Raveaux, Jean-Christophe Burie, Jean-Marc Ogier " A colour document interpretation: Application to ancient cadastral maps " International Conference on Document analysis and Recognition 2007

References - contd

[RRA04] Carol Rids, Corneliu Rusu, and Jaakko Astolu. Contours extraction from a colour coded relief scanned map. International Conference on Image Processing (ICIP), 3:1699 - 1702, Oct 2004.

[Sam84] H. Samet. The quadtree and related hierarchical data structures. ACM Comput. Surveys, 16(2):187 - 260, 1984.

[SG03] Spinello Salvatore and Pierre Pascal Guitton. Contour line recognition from scanned topographic maps. Journal of WSCG, 12(1-3), February 2003.

[SRS03] Biswajit Sarkar, Sanghamitra Roy, and Debranjan Sarkar. Hierarchical representation of digitized curves through dominant point detection. Pattern Recognition Letters, 24:2869 - 2882, 2003.

[SY99] S.C.Huang and Y.N.Sun. Polygonal approximation using genetic algorithms. Pattern Recognition, 32:1409 - 1420, 1999.

[WY00] W.Y.Kim and Y.S.Kim. A region-based shape descriptor using zernike moments. Signal Processing: Image Communication, 16(1-2):95 - 102, 2000.

[XZZ06] Dongjun Xin, Xianzhong Zhou, and Huali Zheng. Contour line extraction from paper-based topographic maps. Journal of Information and Computing Science, 1(5):275 - 283, 2006.

[YLYP98] LeCun Y., Bottou L., Bengio Y., and Haner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278 - 2344, November 1998.

[YMSA08] Y.Ebrahim, M.Ahmed, S.C.Chau, and W. Abdelsalem. Shape representation and description using the hilbert curve. Pattern Recognition Letters, 30(4):345 - 358, 2008.

[ZL04] Dengsheng Zhang and Guojun Lu. Review of shape representation and description techniques. Pattern Recognition, 37:1 - 19, 2004