

Machine vision algorithms on cadaster maps

Sofia Ares Oliveira
Isabella di Lenardo
Frederic Kaplan

Ecole Polytechnique Fédérale de Lausanne
Digital Humanities Laboratory





Al. S. Luca
 Campo S. Luca
 Salvarda
 Campo San Luca
 Salvarda
 Campo S. Luca
 Salvarda

1277
 1275
 1259
 1256
 1257
 1250
 1248
 1247
 1244
 1243
 1242
 1241
 1240
 1239
 1238
 1237
 1236
 1235
 1234
 1233
 1232
 1231
 1230
 1229
 1228
 1227
 1226
 1225
 1224
 1223
 1222
 1221
 1220
 1219
 1218
 1217
 1216
 1215
 1214
 1213
 1212
 1211
 1210
 1209
 1208
 1207
 1206
 1205
 1204
 1203
 1202
 1201
 1200
 1199
 1198
 1197
 1196
 1195
 1194
 1193
 1192
 1191
 1190
 1189
 1188
 1187
 1186
 1185
 1184
 1183
 1182
 1181
 1180
 1179
 1178
 1177
 1176
 1175
 1174
 1173
 1172
 1171
 1170
 1169
 1168
 1167
 1166
 1165
 1164
 1163
 1162
 1161
 1160
 1159
 1158
 1157
 1156
 1155
 1154
 1153
 1152
 1151
 1150
 1149
 1148
 1147
 1146
 1145
 1144
 1143
 1142
 1141
 1140
 1139
 1138
 1137
 1136
 1135
 1134
 1133
 1132
 1131
 1130
 1129
 1128
 1127
 1126
 1125
 1124
 1123
 1122
 1121
 1120
 1119
 1118
 1117
 1116
 1115
 1114
 1113
 1112
 1111
 1110
 1109
 1108
 1107
 1106
 1105
 1104
 1103
 1102
 1101
 1100
 1099
 1098
 1097
 1096
 1095
 1094
 1093
 1092
 1091
 1090
 1089
 1088
 1087
 1086
 1085
 1084
 1083
 1082
 1081
 1080
 1079
 1078
 1077
 1076
 1075
 1074
 1073
 1072
 1071
 1070
 1069
 1068
 1067
 1066
 1065
 1064
 1063
 1062
 1061
 1060
 1059
 1058
 1057
 1056
 1055
 1054
 1053
 1052
 1051
 1050
 1049
 1048
 1047
 1046
 1045
 1044
 1043
 1042
 1041
 1040
 1039
 1038
 1037
 1036
 1035
 1034
 1033
 1032
 1031
 1030
 1029
 1028
 1027
 1026
 1025
 1024
 1023
 1022
 1021
 1020
 1019
 1018
 1017
 1016
 1015
 1014
 1013
 1012
 1011
 1010
 1009
 1008
 1007
 1006
 1005
 1004
 1003
 1002
 1001
 1000
 999
 998
 997
 996
 995
 994
 993
 992
 991
 990
 989
 988
 987
 986
 985
 984
 983
 982
 981
 980
 979
 978
 977
 976
 975
 974
 973
 972
 971
 970
 969
 968
 967
 966
 965
 964
 963
 962
 961
 960
 959
 958
 957
 956
 955
 954
 953
 952
 951
 950
 949
 948
 947
 946
 945
 944
 943
 942
 941
 940
 939
 938
 937
 936
 935
 934
 933
 932
 931
 930
 929
 928
 927
 926
 925
 924
 923
 922
 921
 920
 919
 918
 917
 916
 915
 914
 913
 912
 911
 910
 909
 908
 907
 906
 905
 904
 903
 902
 901
 900
 899
 898
 897
 896
 895
 894
 893
 892
 891
 890
 889
 888
 887
 886
 885
 884
 883
 882
 881
 880
 879
 878
 877
 876
 875
 874
 873
 872
 871
 870
 869
 868
 867
 866
 865
 864
 863
 862
 861
 860
 859
 858
 857
 856
 855
 854
 853
 852
 851
 850
 849
 848
 847
 846
 845
 844
 843
 842
 841
 840
 839
 838
 837
 836
 835
 834
 833
 832
 831
 830
 829
 828
 827
 826
 825
 824
 823
 822
 821
 820
 819
 818
 817
 816
 815
 814
 813
 812
 811
 810
 809
 808
 807
 806
 805
 804
 803
 802
 801
 800
 799
 798
 797
 796
 795
 794
 793
 792
 791
 790
 789
 788
 787
 786
 785
 784
 783
 782
 781
 780
 779
 778
 777
 776
 775
 774
 773
 772
 771
 770
 769
 768
 767
 766
 765
 764
 763
 762
 761
 760
 759
 758
 757
 756
 755
 754
 753
 752
 751
 750
 749
 748
 747
 746
 745
 744
 743
 742
 741
 740
 739
 738
 737
 736
 735
 734
 733
 732
 731
 730
 729
 728
 727
 726
 725
 724
 723
 722
 721
 720
 719
 718
 717
 716
 715
 714
 713
 712
 711
 710
 709
 708
 707
 706
 705
 704
 703
 702
 701
 700
 699
 698
 697
 696
 695
 694
 693
 692
 691
 690
 689
 688
 687
 686
 685
 684
 683
 682
 681
 680
 679
 678
 677
 676
 675
 674
 673
 672
 671
 670
 669
 668
 667
 666
 665
 664
 663
 662
 661
 660
 659
 658
 657
 656
 655
 654
 653
 652
 651
 650
 649
 648
 647
 646
 645
 644
 643
 642
 641
 640
 639
 638
 637
 636
 635
 634
 633
 632
 631
 630
 629
 628
 627
 626
 625
 624
 623
 622
 621
 620
 619
 618
 617
 616
 615
 614
 613
 612
 611
 610
 609
 608
 607
 606
 605
 604
 603
 602
 601
 600
 599
 598
 597
 596
 595
 594
 593
 592
 591
 590
 589
 588
 587
 586
 585
 584
 583
 582
 581
 580
 579
 578
 577
 576
 575
 574
 573
 572
 571
 570
 569
 568
 567
 566
 565
 564
 563
 562
 561
 560
 559
 558
 557
 556
 555
 554
 553
 552
 551
 550
 549
 548
 547
 546
 545
 544
 543
 542
 541
 540
 539
 538
 537
 536
 535
 534
 533
 532
 531
 530
 529
 528
 527
 526
 525
 524
 523
 522
 521
 520
 519
 518
 517
 516
 515
 514
 513
 512
 511
 510
 509
 508
 507
 506
 505
 504
 503
 502
 501
 500
 499
 498
 497
 496
 495
 494
 493
 492
 491
 490
 489
 488
 487
 486
 485
 484
 483
 482
 481
 480
 479
 478
 477
 476
 475
 474
 473
 472
 471
 470
 469
 468
 467
 466
 465
 464
 463
 462
 461
 460
 459
 458
 457
 456
 455
 454
 453
 452
 451
 450
 449
 448
 447
 446
 445
 444
 443
 442
 441
 440
 439
 438
 437
 436
 435
 434
 433
 432
 431
 430
 429
 428
 427
 426
 425
 424
 423
 422
 421
 420
 419
 418
 417
 416
 415
 414
 413
 412
 411
 410
 409
 408
 407
 406
 405
 404
 403
 402
 401
 400
 399
 398
 397
 396
 395
 394
 393
 392
 391
 390
 389
 388
 387
 386
 385
 384
 383
 382
 381
 380
 379
 378
 377
 376
 375
 374
 373
 372
 371
 370
 369
 368
 367
 366
 365
 364
 363
 362
 361
 360
 359
 358
 357
 356
 355
 354
 353
 352
 351
 350
 349
 348
 347
 346
 345
 344
 343
 342
 341
 340
 339
 338
 337
 336
 335
 334
 333
 332
 331
 330
 329
 328
 327
 326
 325
 324
 323
 322
 321
 320
 319
 318
 317
 316
 315
 314
 313
 312
 311
 310
 309
 308
 307
 306
 305
 304
 303
 302
 301
 300
 299
 298
 297
 296
 295
 294
 293
 292
 291
 290
 289
 288
 287
 286
 285
 284
 283
 282
 281
 280
 279
 278
 277
 276
 275
 274
 273
 272
 271
 270
 269
 268
 267
 266
 265
 264
 263
 262
 261
 260
 259
 258
 257
 256
 255
 254
 253
 252
 251
 250
 249
 248
 247
 246
 245
 244
 243
 242
 241
 240
 239
 238
 237
 236
 235
 234
 233
 232
 231
 230
 229
 228
 227
 226
 225
 224
 223
 222
 221
 220
 219
 218
 217
 216
 215
 214
 213
 212
 211
 210
 209
 208
 207
 206
 205
 204
 203
 202
 201
 200
 199
 198
 197
 196
 195
 194
 193
 192
 191
 190
 189
 188
 187
 186
 185
 184
 183
 182
 181
 180
 179
 178
 177
 176
 175
 174
 173
 172
 171
 170
 169
 168
 167
 166
 165
 164
 163
 162
 161
 160
 159
 158
 157
 156
 155
 154
 153
 152
 151
 150
 149
 148
 147
 146
 145
 144
 143
 142
 141
 140
 139
 138
 137
 136
 135
 134
 133
 132
 131
 130
 129
 128
 127
 126
 125
 124
 123
 122
 121
 120
 119
 118
 117
 116
 115
 114
 113
 112
 111
 110
 109
 108
 107
 106
 105
 104
 103
 102
 101
 100
 99
 98
 97
 96
 95
 94
 93
 92
 91
 90
 89
 88
 87
 86
 85
 84
 83
 82
 81
 80
 79
 78
 77
 76
 75
 74
 73
 72
 71
 70
 69
 68
 67
 66
 65
 64
 63
 62
 61
 60
 59
 58
 57
 56
 55
 54
 53
 52
 51
 50
 49
 48
 47
 46
 45
 44
 43
 42
 41
 40
 39
 38
 37
 36
 35
 34
 33
 32
 31
 30
 29
 28
 27
 26
 25
 24
 23
 22
 21
 20
 19
 18
 17
 16
 15
 14
 13
 12
 11
 10
 9
 8
 7
 6
 5
 4
 3
 2
 1
 0

1277
 1275
 1259
 1256
 1257
 1250
 1248
 1247
 1244
 1243
 1242
 1241
 1240
 1239
 1238
 1237
 1236
 1235
 1234
 1233
 1232
 1231
 1230
 1229
 1228
 1227
 1226
 1225
 1224
 1223
 1222
 1221
 1220
 1219
 1218
 1217
 1216
 1215
 1214
 1213
 1212
 1211
 1210
 1209
 1208
 1207
 1206
 1205
 1204
 1203
 1202
 1201
 1200
 1199
 1198
 1197
 1196
 1195
 1194
 1193
 1192
 1191
 1190
 1189
 1188
 1187
 1186
 1185
 1184
 1183
 1182
 1181
 1180
 1179
 1178
 1177
 1176
 1175
 1174
 1173
 1172
 1171
 1170
 1169
 1168
 1167
 1166
 1165
 1164
 1163
 1162
 1161
 1160
 1159
 1158
 1157
 1156
 1155
 1154
 1153
 1152
 1151
 1150
 1149
 1148
 1147
 1146
 1145
 1144
 1143
 1142
 1141
 1140
 1139
 1138
 1137
 1136
 1135
 1134
 1133
 1132
 1131
 1130
 1129
 1128
 1127
 1126
 1125
 1124
 1123
 1122
 1121
 1120
 1119
 1118
 1117
 1116
 1115
 1114
 1113
 1112
 1111
 1110
 1109
 1108
 1107
 1106
 1105
 1104
 1103
 1102
 1101
 1100
 1099
 1098
 1097
 1096
 1095
 1094
 1093
 1092
 1091
 1090
 1089
 1088
 1087
 1086
 1085
 1084
 1083
 1082
 1081
 1080
 1079
 1078
 1077
 1076
 1075
 1074
 1073
 1072
 1071
 1070
 1069
 1068
 1067
 1066
 1065
 1064
 1063
 1062
 1061
 1060
 1059
 1058
 1057



Numeri della Mappa	POSSessori	Denominazione dei Pezzi di terra	QUALITÀ		SUPERFICIE	
			Classe	Pertiche Censuarie	Gen- tesimi	
		Porte di Rialto				
9001	Trimani	30				
9002	Marcello Marino q. ^o Guido	29	C.S.			
9003	Venier Bro' Ferolamo q. ^o Pio: Batta	28	C.S.			
9004	Suddetto	27	C.S.			
9005	Marconi Giuseppe q. ^o Ant. ^o	26	C.S.			
9006	Raspi Fran. ^o q. ^o Pio: Maria	25	C.S.			
9007	Besenti Andrea q. ^o	4885	C.S.			
9008	Morosini Ant. ^o di Vincenzo Dolfin Luigi q. ^o Dolfin Marcant. ^o q. ^o Ferracina Niccolò q. ^o Betrogalli Piu. ^o q. ^o Berganti Piu. ^o q. ^o Dolfin Marianna q. ^o Cognepori indivisi	1580	C.S.			
9009	Maruzzi Costantino q. ^o Canno	4579	C.S.			
9010	Dolfin Leonardo q. ^o Bro'	4578	C.S.			
9011	Venier Ferolamo q. ^o Pio: Batta	4577	C.S.			
9012	Erceli del fu Carlo Emilio Canal q. ^o Ferolamo. Balbi Canal Mattia Ant. ^o Vedo va del fu Cristoforo Cognepori indivisi	4577	C.S.			
9013	Maruzzi Costantino q. ^o Canno	4587	C.S.			
9014	Correr Marcello Maria q. ^o Bro'	4586	C.S.			
9015	Zabia Fran. ^o q. ^o Casto Ant. ^o	4585	C.S.			
9016	Venier Bro' Ferolamo q. ^o Pio: Batta	4584	C.S.			

PLANO DE MADRID.

1879



I. Palacio. II. Universidad. III. Centro. IV. Hospicio. V. Baratas. VI. Congreso. VII. Hospital. VIII. Juicio. IX. Latina. X. Audiencia.



CITY OF MONTREAL
OF THE
SUBURBS



Scale: 1 inch to the mile or 25 miles



Cote St. Louis

Cote St. Catherine

Cote St. Antoine

Coteau St. Pierre

Cote St. Paul

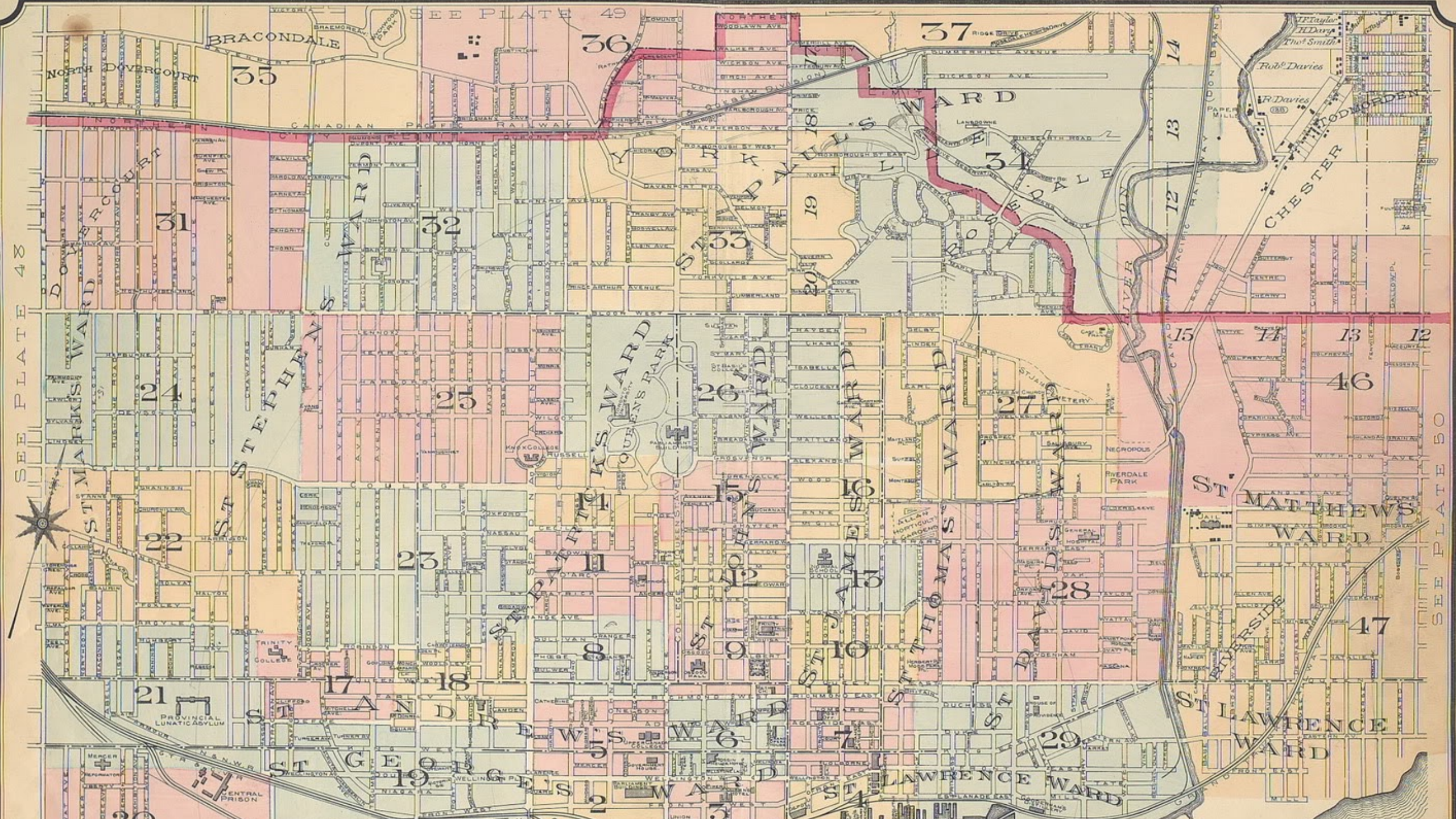
HOCHELAGA P.O.

St. Lawrence
River

POINT ST. CHARLES P.O.

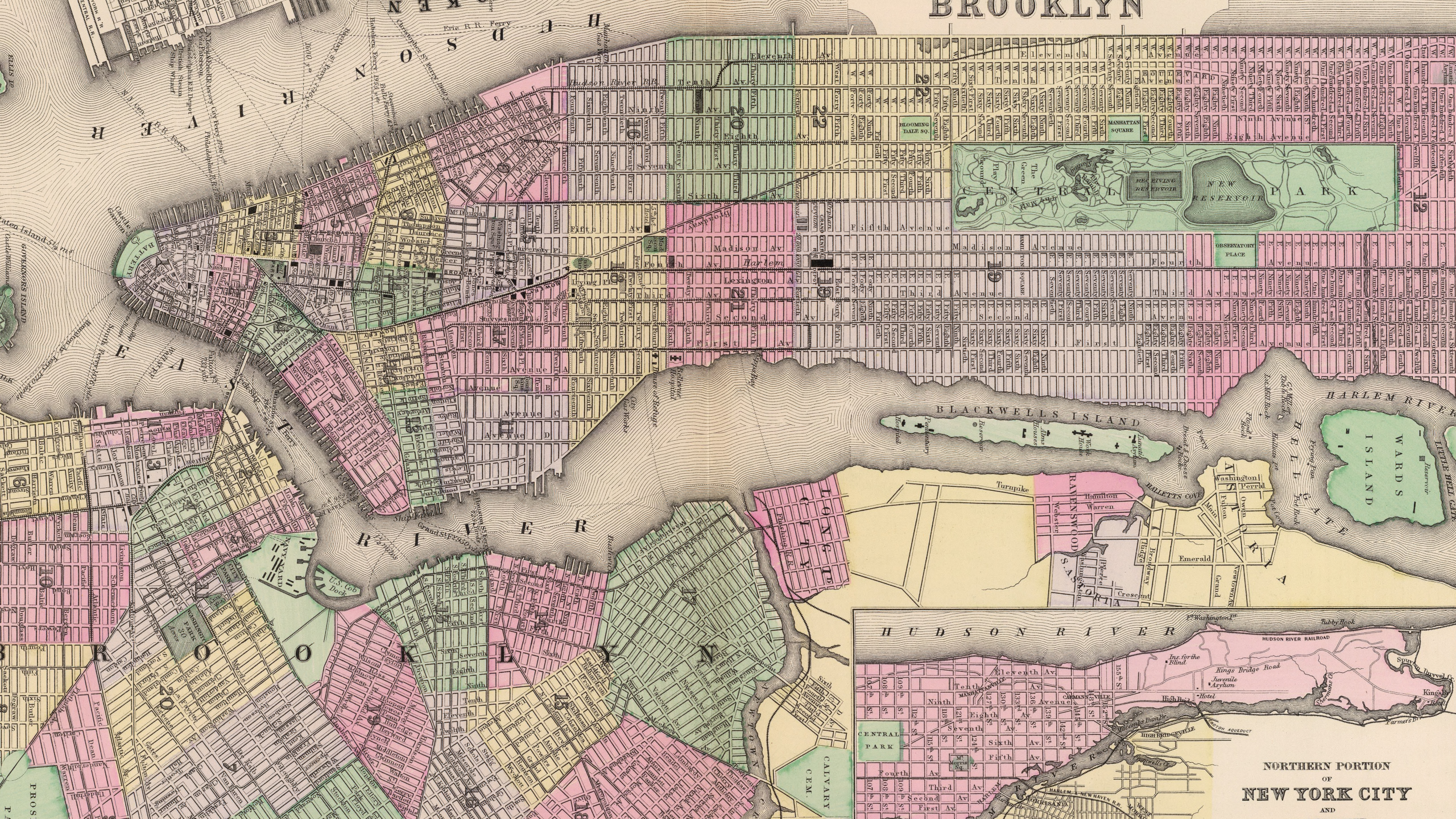
ILE ST. PAUL
OR
NUNS ISLAND

SEE PLATE 48



SEE PLATE 50

BROOKLYN

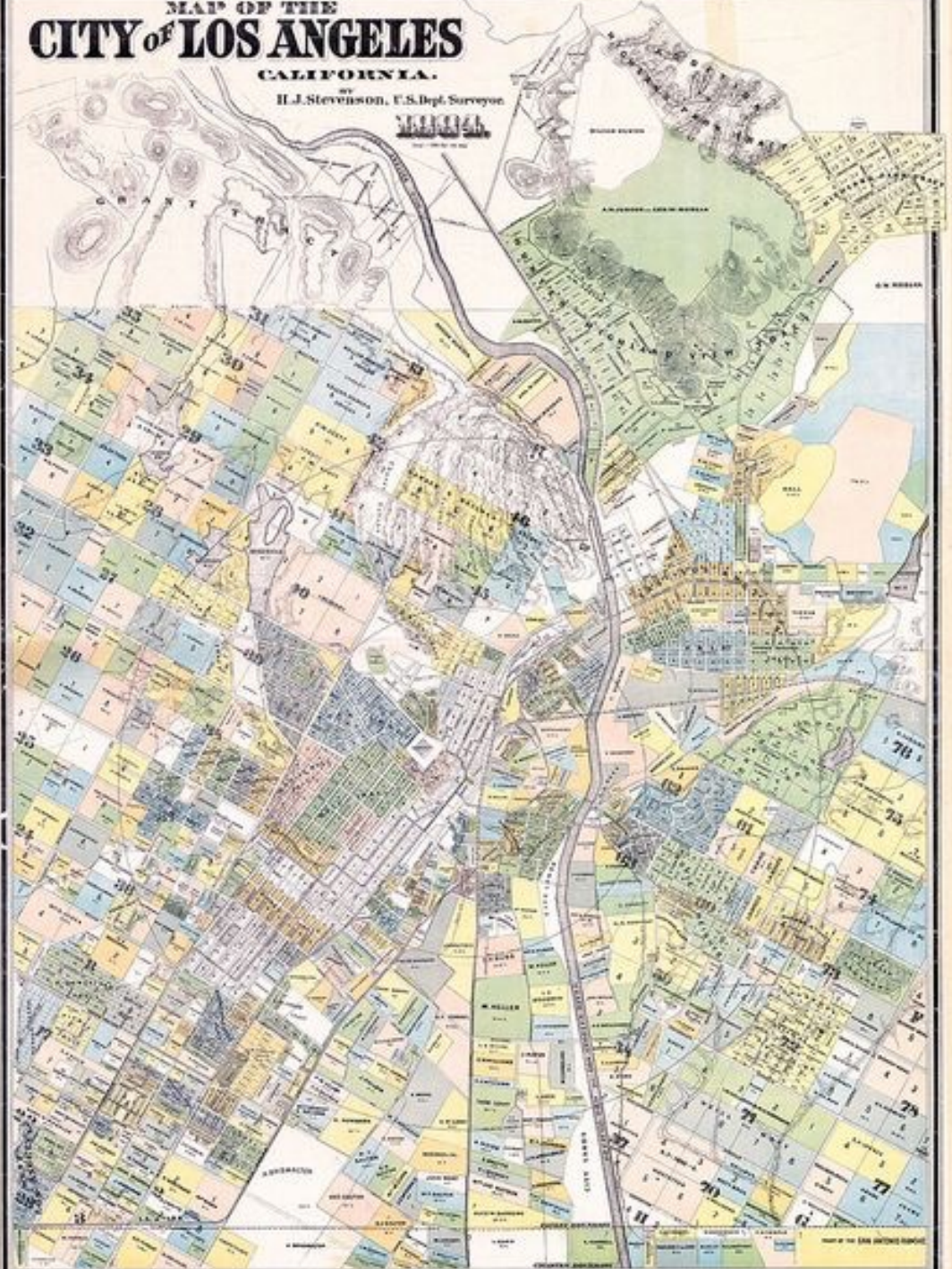


NORTHERN PORTION
OF
NEW YORK CITY
AND

MAP OF THE
CITY OF LOS ANGELES
CALIFORNIA.

H.J. Stevenson, U.S. Dept. Surveyor

MUNICIPAL
MAP



FOR ADJOINING TERRITORY SEE VOLUME I

Scale 100 feet - inch.

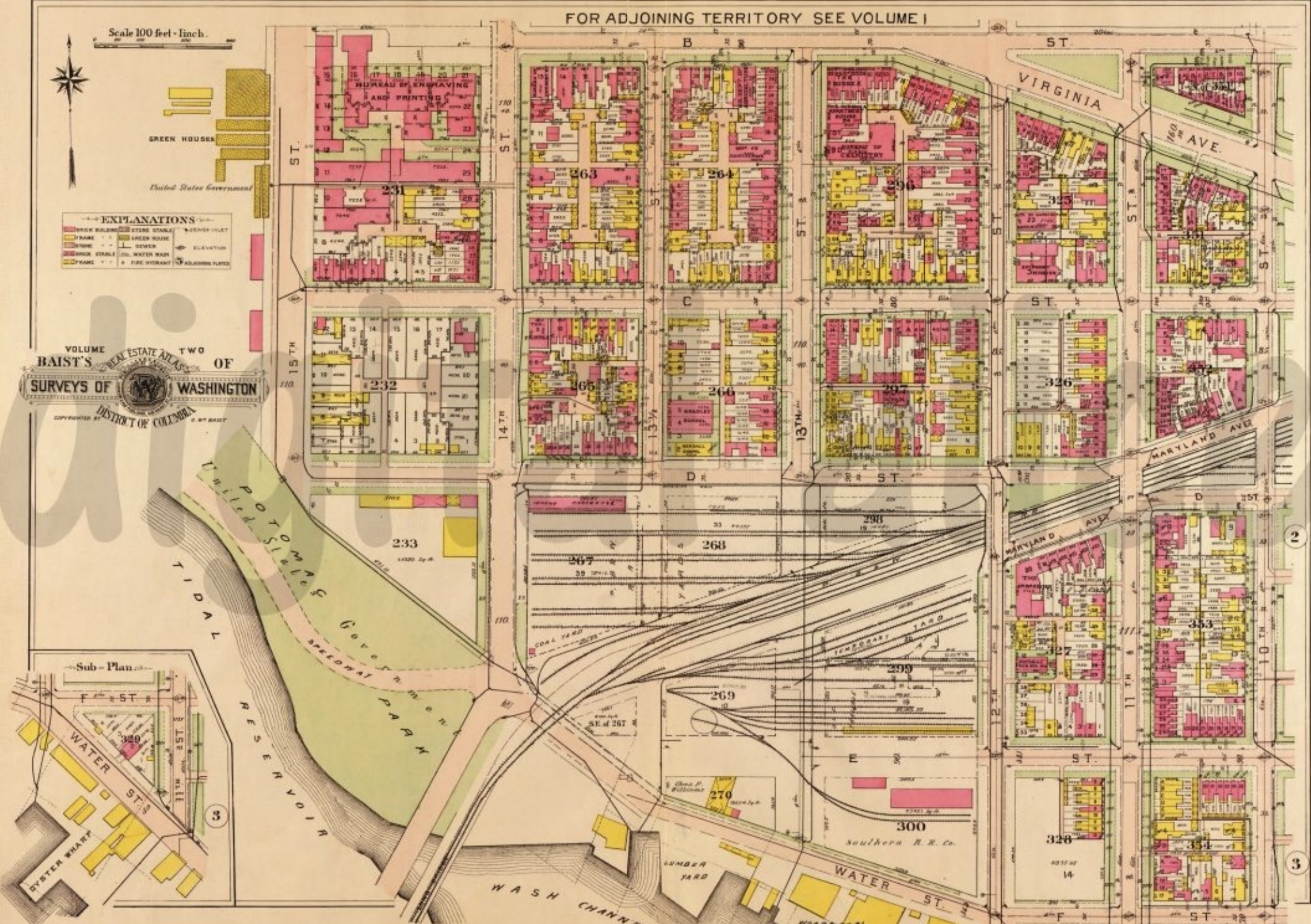


GREEN HOUSES
United States Government

EXPLANATIONS

	BRICK BUILDING		STONE BUILDING		SEWER INLET
	FRAME BUILDING		GREEN HOUSE		SEWER
	BRICK STABLE		WATER MAIN		ELEVATION
	FRAME BUILDING		FIRE HYDRANT		BUILDING PLATES

VOLUME TWO
BAIST'S REAL ESTATE ATLAS OF
SURVEYS OF WASHINGTON
DISTRICT OF COLUMBIA
COPYRIGHTED BY B. BAIST



2

3



1,000 years of archives of the Republic of Venice, in one place



Method





RE
Piazza di S. Giacomo di Prato
I. Euteria
Piazza di S. Giacomo di Prato

AB
M. S. Giovanni Crisostomo
Corte del Teatro
Corte Sabonera
Corte Morosina
Corte di S. Marina
Corte di S. Antonio

AA
Corte di S. Marina
Corte di S. Antonio
Corte di S. Maria Formosa
Corte di S. Antonio

RE
Piazza di S. Giacomo di Prato
Piazza di S. Giacomo di Prato
Piazza di S. Giacomo di Prato
Piazza di S. Giacomo di Prato

YY
S. Marina
Corte di S. Marina
Corte di S. Antonio
Corte di S. Antonio

BB
Corte di S. Marina
Corte di S. Antonio
Corte di S. Antonio
Corte di S. Antonio

II
S. Salvatore
Corte di S. Salvatore
Corte di S. Salvatore
Corte di S. Salvatore

TT
Madonna della Fava
Corte di S. Antonio
Corte di S. Antonio
Corte di S. Antonio

DB
S. Maria Formosa
Corte di S. Antonio
Corte di S. Antonio
Corte di S. Antonio

LL
Corte di S. Salvatore
Corte di S. Salvatore
Corte di S. Salvatore
Corte di S. Salvatore

RR
Corte di S. Antonio
Corte di S. Antonio
Corte di S. Antonio
Corte di S. Antonio

DC
Campelli Quirini
Corte di S. Antonio
Corte di S. Antonio
Corte di S. Antonio

Denoising and ridge detection



Non-Local Means denoising

Ridge detection

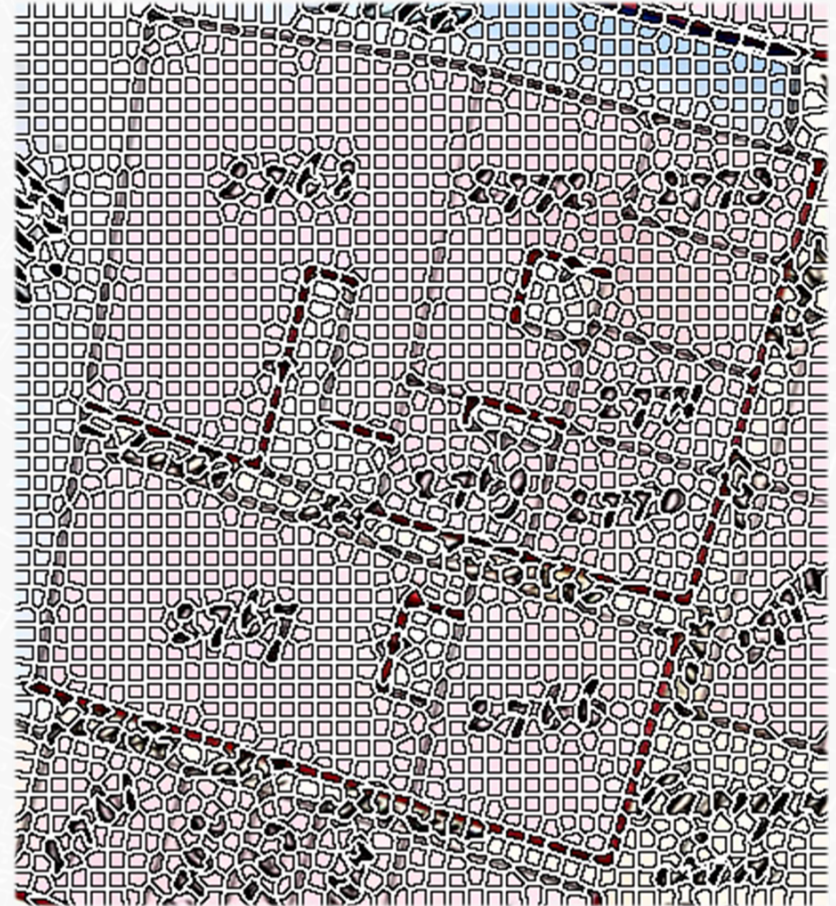


1. Frangi et al., Multiscale vessel enhancement filtering

Simple Linear Iterative Clustering (SLIC superpixels)

Superpixel : cluster of pixels that share similarity and spatial proximity

- Captures image redundancy
- Convenient primitive representation of the image
- Reduces the complexity of subsequent processing tasks

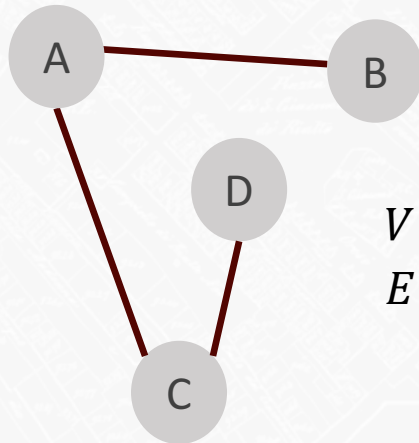


2. Achanta et al., SLIC superpixels compared to state-of-the-art superpixel methods

Graph of superpixels

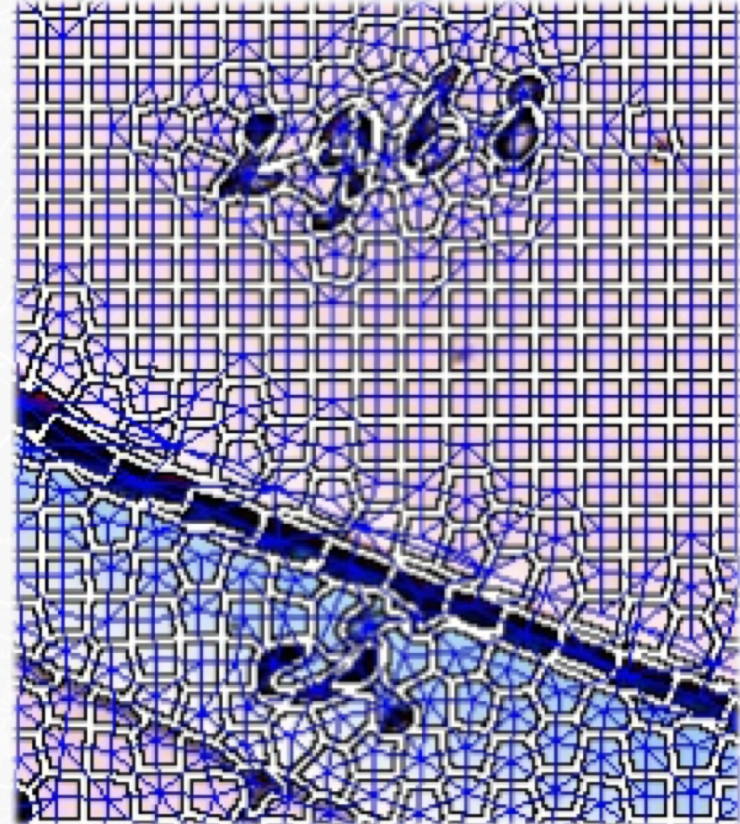
Graph : Set of nodes **V** and edges **E**

- Nodes: superpixels
- Edges: similarity measure between 2 superpixels



$$V = \{A, B, C, D\}$$

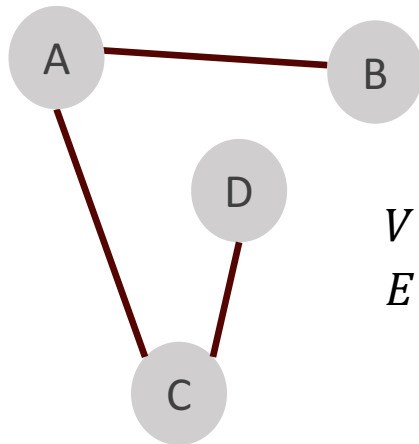
$$E = \{(A, B), (A, C), (C, D)\}$$



Graph of superpixels

Graph : Set of nodes **V** and edges **E**

- Nodes: superpixels
- Edges: similarity measure between 2 superpixels



$$V = \{A, B, C, D\}$$

$$E = \{(A, B), (A, C), (C, D)\}$$

Similarity measure (distance):

- Color distance (CIE2000) based on L,a,b color components

$$D(v_i, v_j) = \Delta E_{2000}(m_i, m_j)$$

v_i, v_j : nodes

ΔE_{2000} : CIE2000 color difference (standard)

m_i, m_j : mean of L,a,b components

Region merging

Group similar superpixels to create bigger regions → Merge nodes whose edges have low values

Merging process

- Minimize intragroup dissimilarity by eliminating edges with high values
- Stopping criterion : homogeneity of the merged region (measured by the 'dispersion' of the edge's values)



Region classification

- Separation of the regions in 3 classes:
 - Text (**T_x**)
 - Contours (**C**)
 - Background (**B**)
- Classifier: Support Vector Machine (SVM)
- Training data:
 - Samples of cadaster maps coming from the venetian cadaster



Parcel extraction : flood fill

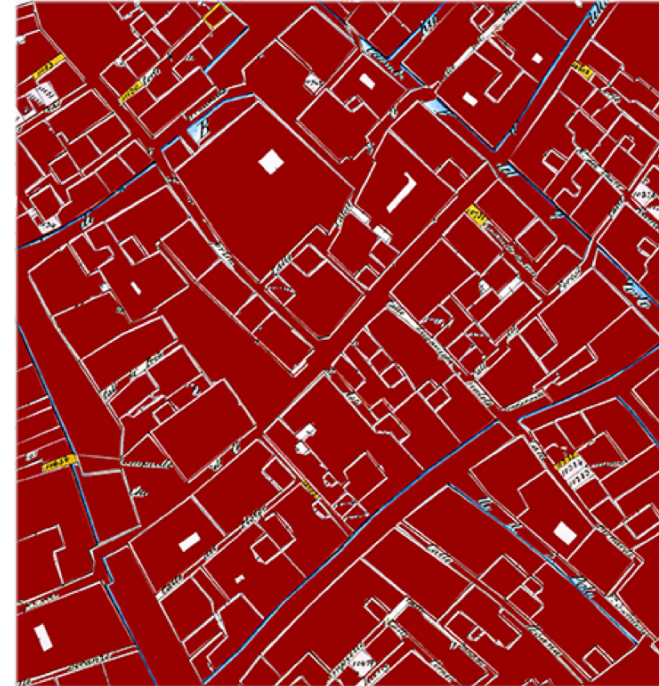
- Regions classified as B are used as seed points and the result of ridge detection filter is used to delimit the regions
- Flood fill :
 - Starting point initialization (seed point)
 - Filling of the region by pixel-to-pixel diffusion
 - Stop the diffusion when all the pixels delimited by the contours have been filled



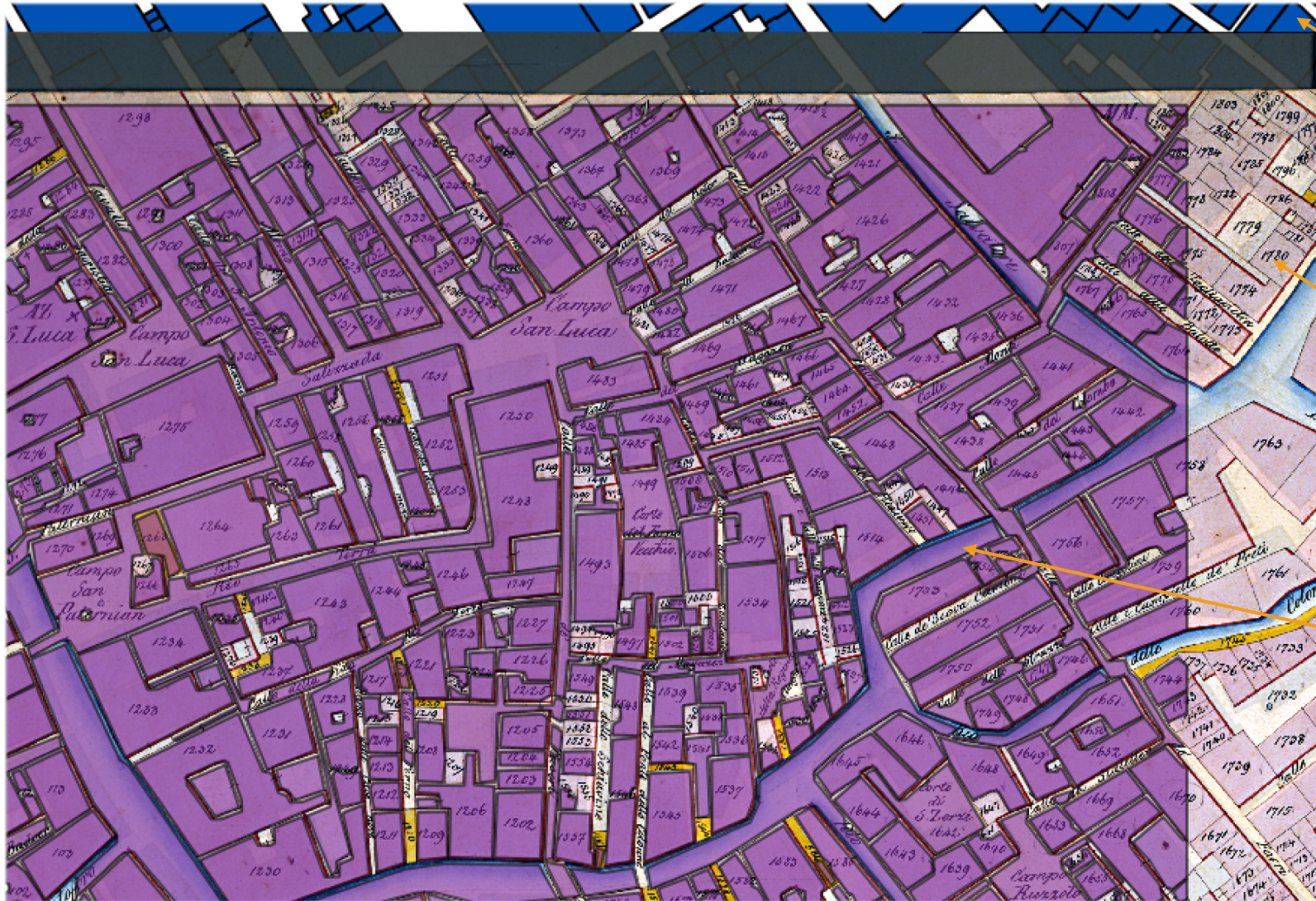
Parcel extraction : polygons and georeferencing

- Approximation of filled regions into polygons
- Corner coordinates according to image referential (x,y) exported into GeoJSON format
- If image is georeferenced → exported polygons can directly be imported into GIS software

```
{"geometry": {"coordinates": [[[2311867.7352005076,  
5035105.466485897], [2311864.413514898,  
5035097.681285249], [2311865.451541651,  
5035096.850863847], [2311865.451541651,  
5035095.397626393], [2311863.7906988463,  
5035091.3493220555], [2311870.2264647153,  
5035089.4808739005], [2311874.3785717273,  
5035102.663813664]]]], "type": "Polygon"}  
  , ... }
```



Georeferenced parcels



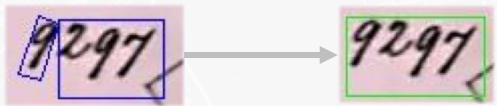
Current parcels

Cadaster map image
(georeferenced)

Extracted parcels
(from 1808)

Parcel's label extraction

- Locating:
 - Locate regions labelled as τ_x
 - Group digits of same identifiers
- Orientation correction using Principal Component Analysis (PCA)



- Extraction of boxes containing labels

Label transcription with CRNN

- Convolutional Recurrent Neural Network (CRNN)
 - Convolutional layers : compact encoding of image information
 - Recurrent layers (LSTM) : handle sequences of arbitrary length
- Training data :
 - Synthetic data generated from MNIST dataset (100K)
 - Handwritten numbers extracted from venetian archives (~ 30K)

2338. 1242817 1157
46995 7383 9329519

5599	Pred: 5599
3457	Pred: 6457
2633	Pred: 2633
3955	Pred: 5935
6682	Pred: 6682
2704	Pred: 2704
2686	Pred: 2686
6681	Pred: 66817

3. Shi et al, An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition

Results

Parcel extraction results



Parcel extraction results

IoU threshold	0.6	0.7	0.8
Parcel recall	0.77 (623)	0.76 (616)	0.72 (583)
Parcel precision	0.55	0.54	0.51
Extracted parcels	1144		
Ground truth (parcels with labels)	810		

$$recall = \frac{\text{true positives}}{\text{total ground truth}} \in [0, 1]$$

$$precision = \frac{\text{true positives}}{\text{total retrieved}} \in [0, 1]$$

Label extraction and transcription results



3329

3329 (score : 7.35)

3457

6457 (score : 1.37)

3322

3322 (score : 5.19)

3290

3290 (score : 10.52)

3326

332 (score : 1.48)

3342

3342 (score : 5.77)

3345-

3345 (score : 10.29)

3333

3333 (score : 3.97)

3341

334 (score : 4.74)

3441

3341 (score : 10.76)

3368

336 (score : 2.70)

3321

3321 (score : 1.15)

3349

3349 (score : 3.77)

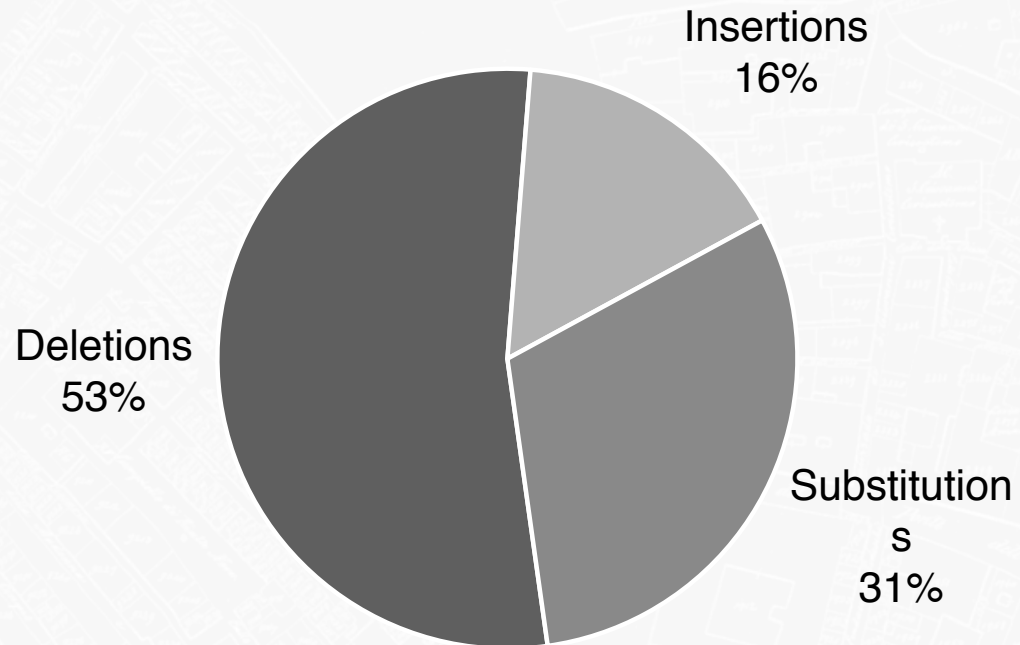
Label extraction and transcription results

Label locating	Inter : 0.8	Inter: 0.9	IoU : 0.5
Recall	0.81 (596)	0.77 (568)	0.36 (266)
Precision	0.59	0.57	0.26
Extracted labels	1004		
Ground truth (labels)	736		

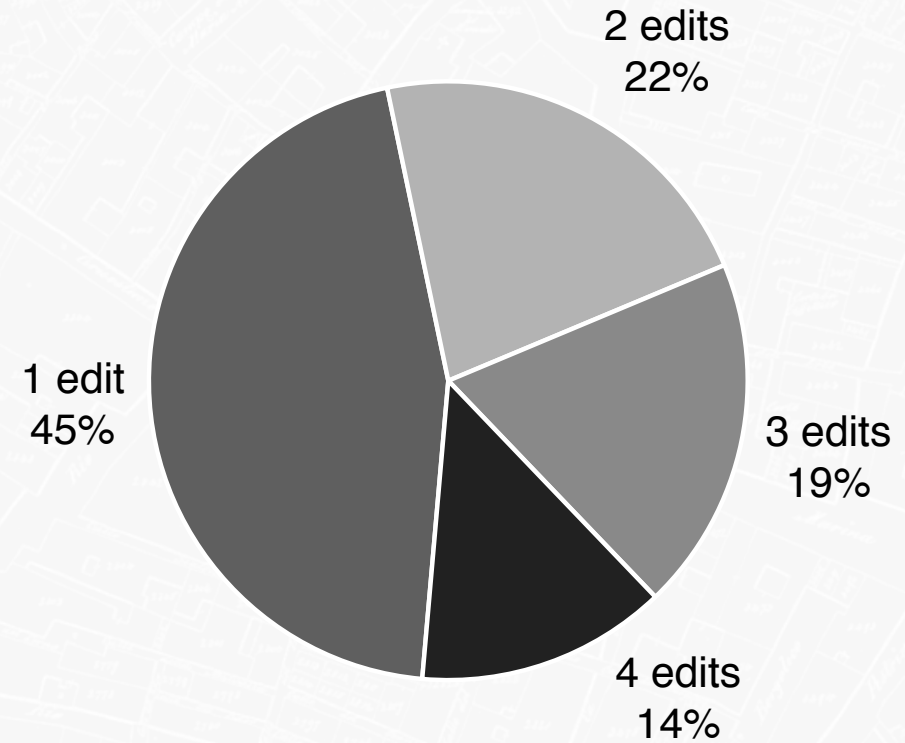
Label transcription	Inter : 0.8	Inter: 0.9	IoU : 0.5
Recall (correctly transcribed)	0.50 (367)	0.48 (356)	0.28 (208)
Precision	0.62	0.63	0.78
Character Error Rate (CER)	0.20	0.19	0.10
Ground truth (labels)	736		

Label transcription: what are the errors ?

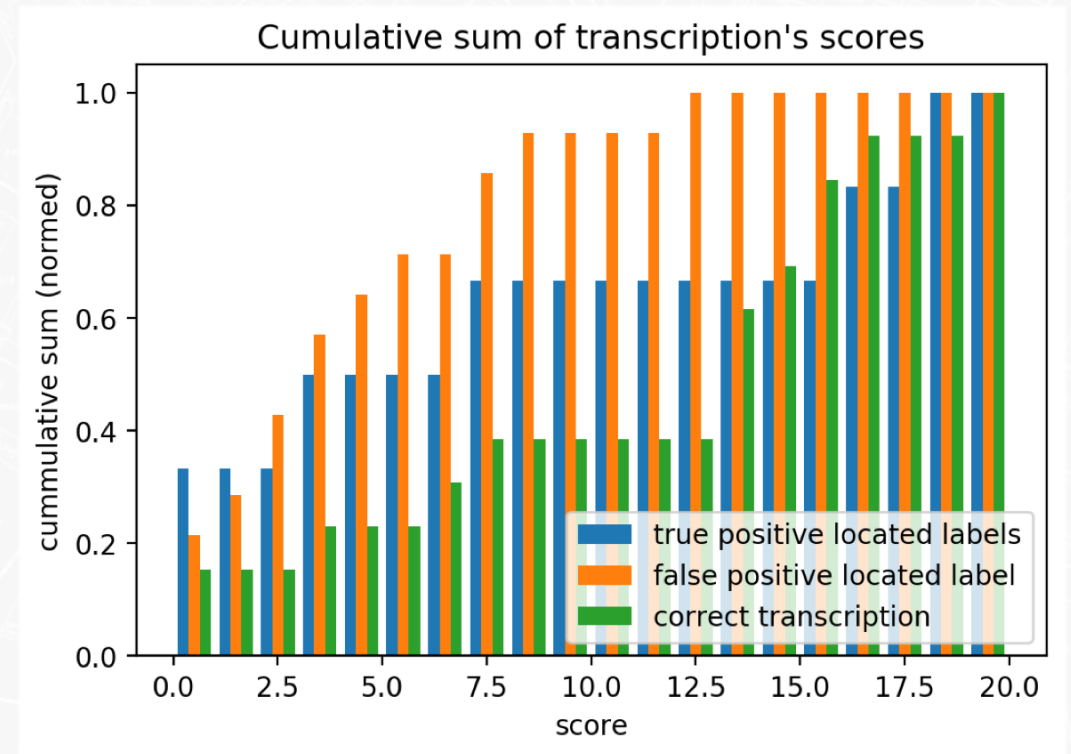
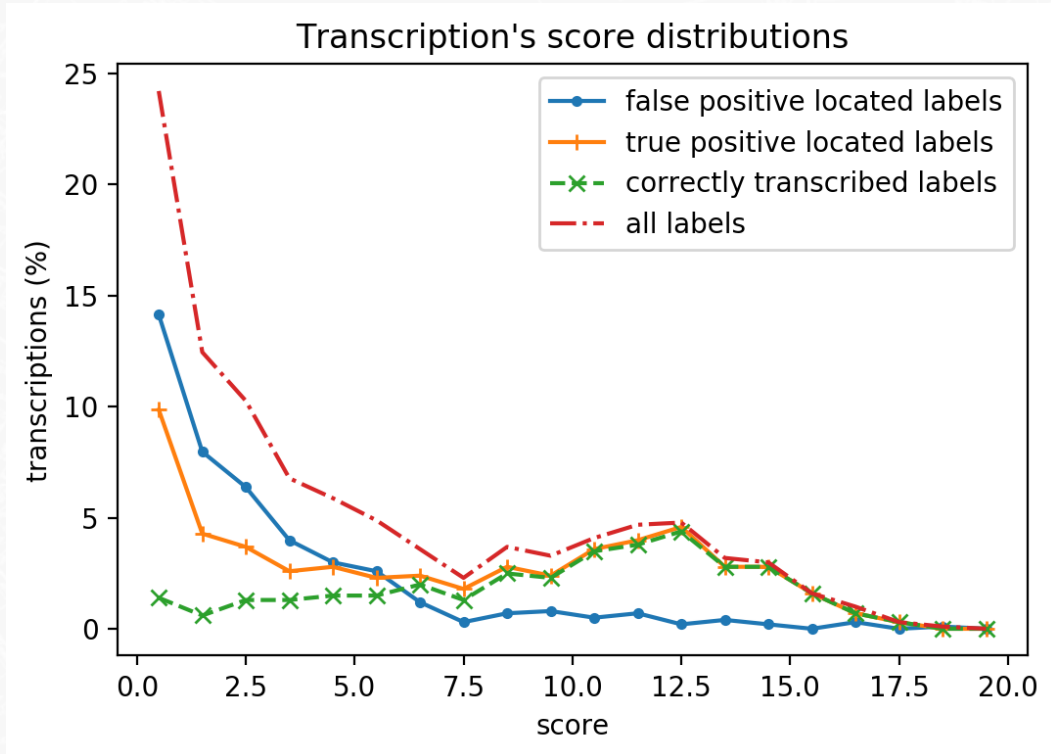
Type of errors



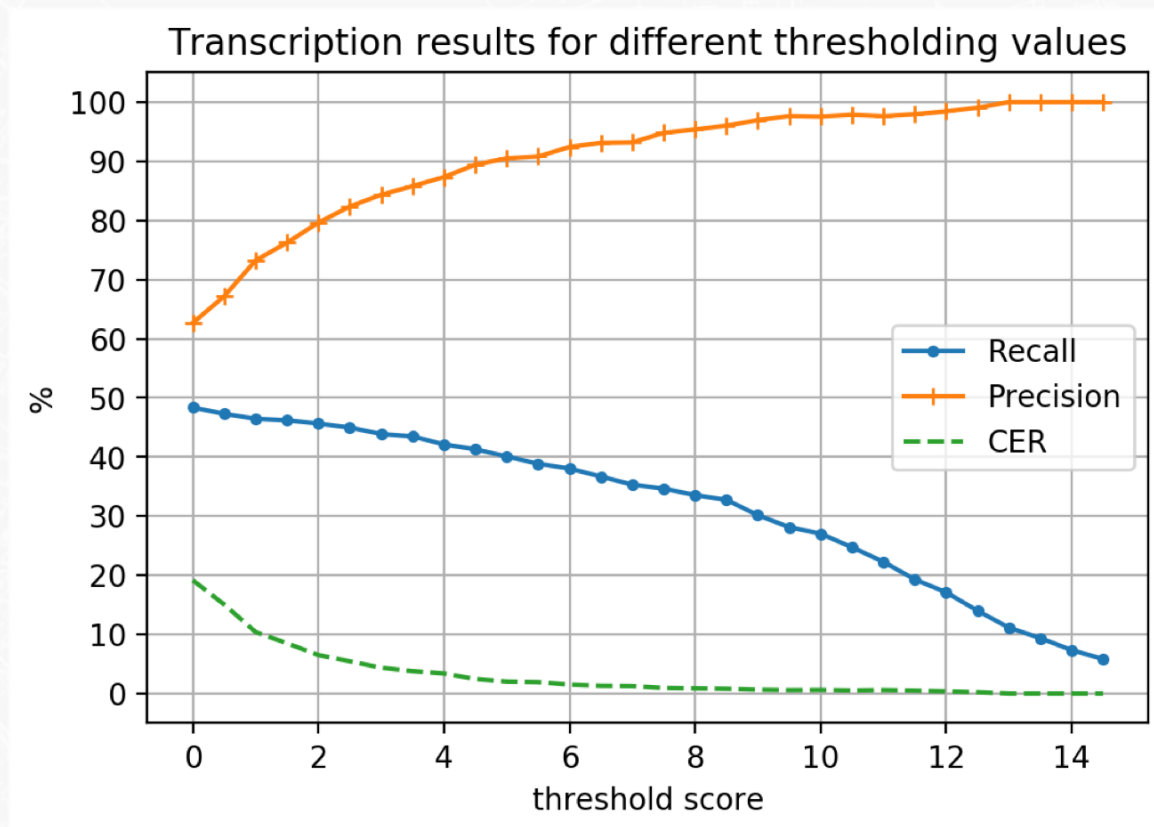
Edit distance



False positive labels and transcription errors : how to detect them ?



Decreasing false positive rate using scores



Threshold on score	0.0	3.0
Recall transcription	0.50 (367)	0.45 (334)
Precision transcription	0.62	0.84
CER	0.20	0.05
Total transcriptions	596	399



Full working pipeline to process cadaster images

50 % labels correctly transcribed

72 % parcels correctly extracted

Precision can be increased using score's information

Contact information

Digital Humanities Laboratory dhlab.epfl.ch

Demo version of the code github.com/dhlab-epfl/cadasters

Sofia Ares Oliveira sofia.oliveiraares@epfl.ch

Isabella di Lenardo isabella.dilenardo@epfl.ch

Frederic Kaplan frederic.kaplan@epfl.ch

References

1. Frangi, A. F., Niessen, W. J., Vincken, K. L., & Viergever, M. A. (1998, October). Multiscale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 130-137). Springer Berlin Heidelberg.
2. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11), 2274-2282.
3. Shi, B., Bai, X., & Yao, C. (2016). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE transactions on pattern analysis and machine intelligence*.