

# Index

## A

- Abolghasemi, V., 174
- Advertising systems
  - next-generation video ad system, 190
  - online video advertising, 188–189
  - VideoSense, 189–190
- Agnihotri, L., 196
- Ahmadyfard, A., 174
- Allier, B., 134
- Anagnostopoulos, C.N.E., 173
- Anthimpoulos, M., 72
- Aradhye, H., 242
- Arithmetic mean filter, 52
- Automatic speech recognition (ASR), 230–231

## B

- Bahu, B.R., 116, 120
- Bangla and Devanagari text, 197
- Barnard, K., 228
- Baseball videos. *See* Sport video analysis systems
- Bayesian classifier
  - BGM, 77
  - language-independent multi-oriented text detection, 200
  - Laplacian-Sobel product process, 74
  - text candidates, 73
  - text enhancement, 73–75
  - true text pixel classification, 75–76
- Binarization method, 148
  - Canny-based binarization method, 159
  - character segmentation and word, 148
  - post-processing techniques (*see* Post-processing techniques)
  - preprocessing techniques, 172–173

- Blei, D.M., 228
- Boundary growing method (BGM), 77, 202–203
- Bresenham's line drawing algorithm, 149
- Broadcast video content analysis
  - classical text-based information retrieval system, 230
  - Google *Tesseract*, 230
  - MAP, 231
  - MED, 231
  - story boundary detection, 231–232
  - text-based person search system, 230
  - video OCR/ASR, 230–231
- Busch, A., 208

## C

- Campos, T.E.D., 116, 120
- Caner, H., 179
- Canny-based binarization method, 159
- Caption text
  - connected component-based methods
    - color reduction, 63
    - input image, 64
    - Laplacian operation, 65
    - segmentation, 66–67
    - simple and complex, 66
    - text candidate detection, 65
  - edge-based methods
    - arithmetic mean filter, 52
    - block image, 52, 53
    - candidate text block, 56
    - Canny operator, 52
    - edge filter, 51
    - feature selection, 56
    - grayscale image, 52, 53

- Caption text (*cont.*)
      - k-means clustering algorithm, 52
      - median filter, 52
      - RGB color models, 51
      - segmented region, 56
      - Sobel edge maps, 52
      - vertical and horizontal bars, 56
      - weak edges, 55
    - Fourier and Laplacian filtering
      - final text detection result, 71
      - ideal low-pass filter, 69–70
      - maximum difference, 70, 71
      - text and non-text regions, 68
      - text cluster, 71
    - graphics station logo, 49
    - location and size, 49
    - machine learning-based methods
      - Bayesian classifier (*see* Bayesian classifier)
      - neural network, 72–73
      - support vector machine, 72
    - texture-based methods
      - chip generation stages, 57
      - false-positives filtering, 61–62
      - Haar wavelet decomposition, 57
      - license plate localization, 57
      - nonlinear transformation, 57
      - non-text regions, 58
      - regions of interests, 58
      - support vector machines, 58
      - text regions, 58
      - texture features, 59–61
      - texture segmentation, 57
    - video topics, 49
  - Chang, S.F., 172, 185
  - Character extraction, 153
  - Character recognition, 146
    - Canny-based binarization method, 159
    - hierarchical classification
      - binary tree, 161–164
      - features for, 160–161
      - sample size, 160
      - uppercase letters, 161
    - holistic approach, 159
    - LPR systems, 177–180
    - Monte Carlo sampling, 159
    - OCR engines, 160
    - structural features, 164–166
  - Character reconstruction
    - border gap filling, 140–141
    - connected component analysis, 134
    - contour reconstruction, 134
    - horizontal and vertical gap filling, 138–139
    - horizontal and vertical medial axes, 136–138
    - large gap filling, 139–140
    - medial pixels, 134–135
    - natural scene images, 133
    - ring radius transform, 134–136
    - small gap filling, 141–142
    - video character, 133
  - Character segmentation
    - Chinese text detection, 196
    - LPR systems, 177–178
    - without word segmentation
      - candidate cut pixels, 156–157
      - false-positive elimination, 158–159
      - GVF-based method, 155
      - minimum-cost pathfinding problem, 157–158
  - and word
    - binarization method, 148
    - Bresenham's line drawing algorithm, 149
    - character extraction, 153
    - character gap identification, 154
    - error rates, 148
    - Fourier-and moments-based method, 148–149
    - Fourier-moments features, 150–152
    - Monte Carlo sampling, 148
    - multi-oriented text detection, 147
    - scene text detection, 147
    - semantic gap, 147
    - sports events analysis, 147
    - text region extraction methods, 147–148
    - video events analysis, 147
    - word extraction, 150, 152
- Chaudhuri, B.B., 209
- Chen, D., 52, 148, 159
- Chen, K.-T., 230
- Chinese and Japanese Kanji characters, 200
- Chinese text, 196
- Chih-Yi, C., 185
- Classical text-based information retrieval system, 230
- Color homogeneity, 91
- Cong, Y., 110
- Connected components (CCs)
  - caption text
    - color reduction, 63
    - input image, 64
    - Laplacian operation, 65
    - segmentation, 66–67
    - simple and complex, 66

- text candidate detection, 65
- post-processing techniques, 134
- preprocessing techniques, 29
- scene text detection, 91

Cross-media analysis, 229

Cross-media correlation modeling approach, 228–229

**D**

David, A., 228

Devanagari scripts, 197

Dimitrova, N., 196

Dori, D., 248

**E**

Edge-based methods

- caption text
  - arithmetic mean filter, 52
  - block image, 52, 53
  - candidate text block, 56
  - Canny operator, 52
  - edge filter, 51
  - feature selection, 56
  - grayscale image, 52, 53
  - k-means clustering algorithm, 52
  - median filter, 52
  - RGB color models, 51
  - segmented region, 56
  - Sobel edge maps, 52
  - vertical and horizontal bars, 56
  - weak edges, 55
- preprocessing techniques, 35, 36

Emptoz, H., 134

English text, 196

Epshtein, B., 91

**F**

Farsi/Arabic texts, 196

Fourier-and moments-based method, 148–149

Fourier-Laplacian transformations, 94–95

Fusing attention models, 236–237

**G**

Garcia, C., 128

Ghosh, D., 209

González, A., 115

Google Map, 182

Google *Tesseract*, 230

Głowacz, A., 43

Graphics station logo, 49

**Graphics text**

- lyrics analysis, 232–233
- scene text detection, 81
- video text detection, 13

Gray level co-occurrence matrix (GLCM), 31–32

**H**

Harel, J., 84

Hasan, Y.M.Y., 27

Hidden Markov model, 185

High-resolution horizontal text camera images (HHCI), 205, 207

High-resolution multi-oriented camera images (HMCI), 205

Histogram-based methods, 35

Histogram of Oriented Gradients (HOGs), 98

Horn-Schunck (HS) method, 41, 42

Hua, X., 248, 250, 251

Huang, Y.-P., 175

**I**

ICDAR 2003 Robust Reading dataset, 116–118, 252

Image engineering (IE), 34

Image segmentation

- preprocessing techniques
  - bottom-up segmentation, 34
  - clustering procedure, 38
  - edge-based methods, 35, 36
  - generalized eigenvalue system, 37
  - graph-based segmentation, 37
  - grayscale image, 35
  - Histogram-based methods, 35
  - mean shift segmentation, 39
  - normalized cut criteria, 36, 37
  - role of, 34
  - shift-based method, 38
  - top-down segmentation, 34
  - video text detection, 11

Indian scripts, 197

International Conference on Document Analysis and Recognition (ICDAR), 249

Iria, J., 228

Itti, L., 86

**J**

Jae-Chang, S., 27

Jain, A.K., 28, 58, 64

Japanese Kanji characters, 200

- Jin, L., 176  
 Jin, Y.K., 240  
 Jolion, J.-M., 159, 248  
 Jordan, M.I., 228  
 Jung-Jin, L., 105  
 Jung, K., 58, 72
- K**
- Kai, W., 113  
 Kanade, T., 51  
 Kankanhalli, A., 63  
 Karam, L.J., 27  
 Karaoglu, S., 87, 226  
 Karaoke music retrieval system, 233  
 Karaoke music videos, 232  
 Karatzas, D., 119  
 Kasturi, R., 250, 252  
 Kim, H.K., 13, 28, 64  
 Kim, K.I., 30, 58  
 Kopf, S., 147  
 Korean texts, 196  
 Kunishige, Y., 102
- L**
- Landais, R., 249  
 Language-dependent text detection  
   Bangla and Devanagari text, 197  
   Chinese text, 196  
   English text, 196  
   Farsi/Arabic texts, 196  
   headline-based method, 197–199  
   Indian scripts, 197  
   Korean texts, 196  
   SVM classifier, 196  
 Language-independent multi-oriented text  
   detection  
     Bayesian classifier, 200  
     boundary growing (BC) method, 202–203  
     Chinese and Japanese Kanji characters, 200  
     English text lines, 200  
     HHCI, 205, 207  
     HMCI, 205  
     ICDAR-2003 competition data, 205, 209  
     Laplacian/Fourier techniques, 200  
     LMCI, 205, 208  
     LMMI, 205, 206  
     non-straight text lines, 200  
     run lengths, 201–202  
     Sobel/Laplacian methods, 200  
     zero crossing method, 204  
 Laplacian filter, 25  
 Laplacian-Sobel product process, 74
- Lee, C.M., 63  
 License plate recognition (LPR) systems  
   annual inspection status, 170–171  
   block diagram, 171  
   character recognition, 177–180  
   character segmentation, 177–178  
   images capture, 171–172  
   license plate detection, 175–176  
   preprocessing techniques  
     binarization, 172–173  
     images quality, 174–175  
     shadow/uneven illumination, 173–174  
     video frames, 171  
     skew correction, 176–177  
 Lien, C.-C., 185  
 Lienhart, R., 13, 64, 249, 251  
 Li, H., 58  
 Li, H.P., 252  
 Line drawings, 9  
 Liu, W., 248  
 Liu, Y., 196  
 Li, X., 72  
 Lombardi, L., 182  
 Low-resolution mobile camera images  
   (LMCI), 205, 208  
 Low-resolution multi-oriented mobile camera  
   images (LMMI), 205, 206  
 Lucas-Kanade method, 43  
 Lu, T., 229, 232  
 Lyrics analysis  
   color changes, 233–234  
   graphics text, 232–233  
   karaoke music retrieval system, 233  
   karaoke music videos, 232  
   lyrics text bitmap images, 235–236  
   lyrics text detection technique, 233  
   MPEG-1 NTSC format, 235  
   music structure analysis, 233  
   techniques, 234–235  
   text boxes, 233–234  
   text lines, 233–234  
   text objects, 233–234  
 Lyu, M.R., 196
- M**
- Magalhaes, J., 228  
 Mao, W., 57  
 Marmo, R., 182  
 Matas, J., 170  
 Maximally stable extremal regions  
   (MSERS), 95  
 Maximum gradient difference (MGD),  
   29–30

- Ma, Y.-F., 236  
 Mean average precision (MAP), 231  
 Mehta, A., 189  
 Mei, T., 189  
 Mel frequency cepstrum coefficient (MFCC), 224  
 Metacafe, 188  
 Minimum edit distance (MED), 231  
 Mishra, A., 111  
 Moradi, M., 196  
 Motion analysis, 19–20  
   benefits, 40, 41  
   character recovery, 40  
   Horn-Schunck (HS) method, 41, 42  
   Lucas-Kanade method, 43  
   optical flow, 41–42  
   optical flow computation methods, 44, 45  
   overlapped word elimination, 40  
   text integration, 40  
   text rotation, 40  
   text tracking, 40  
 MSRA Text Detection 500 Database (MSRA-TD500), 249  
 Multimodal video analysis  
   auditory modality, 221  
   broadcast videos (*see* Broadcast video content analysis)  
   class building, 226–227  
   cross-media analysis, 229  
   cross-media correlation modeling approach, 228–229  
   early fusion, 224–225  
   heterogeneous multimodal features, 222–223  
   IMET dataset, 226  
   late fusion, 224–226  
   lyrics analysis (*see* Lyrics analysis)  
   natural language processing, 222  
   pyramidal theory, 222  
   semantic annotation approach, 228  
   semantic gap, 222  
   SVM model, 226  
   text modality, 223  
   textual modality, 221  
   video retrieval system, 222  
   video summarization (*see* Video summarization)  
   visual and auditory modalities, 223–224  
   visual feature descriptors, 224  
   visual modality, 221  
   web video category/search (*see* Web video category/search)  
 Music structure analysis, 233
- N**  
 Natural Environment OCR (NEOCR) dataset, 252  
 Natural scene images, 133  
 Natural scene text detection methods. *See* Scene text detection  
 Navigation assistant systems  
   digital maps, 182  
   GPS-based navigation system, 183  
   milepost signs, 182  
   multi-scale text detection algorithm, 182  
   neural network-based automated system, 182  
   road signs, 181  
   text-based navigation assistant system, 181  
 Netzer, Y., 120  
 Neumann, L., 170  
 Neural network-based automated system, 182  
 Ntirogiannis, K., 128
- O**  
 Odobez, J., 148, 159  
 Ohya, J., 63  
 Online advertisement, 169–170  
 Optical character recognition (OCR)  
   advantages, 145  
   character recognition, 160  
   Dijkstra’s algorithm, 147  
   engines, 146  
   NEOCR dataset, 252  
   path’s cost, 147  
   projection profile analysis, 146–147  
   scene text detection, 81  
   script identification, 207  
   text line binarization, 127  
   video labeling, 146  
   video text detection, 8–9
- P**  
 Palma, D., 39  
 Pan, J., 97  
 Park, S.H., 57  
 Part-of-speech (POS) tagger, 238  
 Pazio, M., 183  
 Performance evaluation  
   benchmark databases, 248–249  
   matching detected text boxes, 249–250  
   pattern recognition systems/algorithms, 247  
   performance metrics, 250–252  
   protocols, 248  
   vs. video text recognition, 252–253  
 Phan, T.Q., 210

- Poignant, J., 229, 230
- Post-processing techniques
- character reconstruction
    - border gap filling, 140–141
    - connected component analysis, 134
    - contour reconstruction, 134
    - horizontal and vertical gap filling, 138–139
    - horizontal and vertical medial axes, 136–138
    - large gap filling, 139–140
    - medial pixels, 134–135
    - natural scene images, 133
    - ring radius transform, 134–136
    - small gap filling, 141–142
    - video character, 133
  - text line binarization
    - foreground and background separation, 132
    - low contrast and complex background, 127
    - OCR technology, 127
    - smoothing, 131–132
    - text candidates, 129–131
    - wavelet-gradient-fusion method, 128–129
- Preprocessing techniques
- color-based preprocessing
    - color clustering, 29
    - color transformations, 26, 27
    - connected components, 29
    - grayscale image, 27, 28
    - multicolor image, 29
    - RGB image, 26
  - contrast analysis, 20
  - image cropping and local operators, 21–22
  - image enhancement techniques, 20
  - image segmentation
    - bottom-up segmentation, 34
    - clustering procedure, 38
    - edge-based methods, 35, 36
    - generalized eigenvalue system, 37
    - graph-based segmentation, 37
    - grayscale image, 35
    - histogram-based methods, 35
    - mean shift segmentation, 39
    - normalized cut criteria, 36, 37
    - role of, 34
    - shift-based method, 38
    - top-down segmentation, 34
  - morphology operators, 25–26
  - motion analysis, 19–20
    - benefits, 40, 41
    - character recovery, 40
    - Horn-Schunck (HS) method, 41, 42
    - Lucas-Kanade method, 43
    - optical flow, 41–42
    - optical flow computation methods, 44, 45
    - overlapped word elimination, 40
    - text integration, 40
    - text rotation, 40
    - text tracking, 40
  - neighborhood operators
    - average filter, 22
    - image sharpening, 23
    - Laplacian operator, 24
    - linear filter, 21–22
    - median filter, 22
    - Roberts operator, 23
    - Sobel operator, 23, 24
  - texture analysis
    - Gabor filter, 32–33
    - gray level co-occurrence matrix, 31–32
    - intensity variations, 30
    - maximum gradient difference, 29–30
    - spatial variances, 29
    - wavelet transform techniques, 30–31
- Probabilistic latent semantic analysis (PLSA), 222
- Projection profile analysis, 146–147
- Q**
- Qian, X., 33
- R**
- Roberts filter, 24
- Roberts, L.G., 24
- S**
- Saidane, Z., 128
- Scene text detection
- bottom-up approach
    - color homogeneity, 91
    - connected component-based methods, 91
    - Fourier-Laplacian transformations, 94–95
    - maximally stable extremal regions, 95
    - stroke width transform, 91–94
    - text detection rate and low false alarm rate, 95
  - Campos and Bahu's dataset, 121
  - character/text recognition
    - characteristics, 113

- direction histogram, 115
  - KNN approach, 115
  - SVM classifier, 115
  - word detection, 113–114
- computer vision community, 81
- Google Street View, 81
- graphics text, 81
- Harel's GBVS saliency map, 86
- hybrid approach
  - bottom-up grouping procedure, 110, 111
  - candidate linking, 111
  - chain analysis, 111
  - component analysis, 110
  - component extraction, 110
  - top-down pruning procedure, 110, 111
- ICDAR dataset, 116
  - outdoor and indoor scenes, 117
  - real scenes and born-digital images, 117
  - TrialTrain and TrialTest, 117
  - web-based tagging system, 116–117
  - XML fragment, 118
- OCR techniques, 81
- statistical and machine learning approach
  - Bayesian classifier, 106
  - character feature, 102–103
  - city scenes, 103
  - false detections, 103
  - Gentle AdaBoost, 105
  - Hough transform, 101
  - Hough vote images, 101
  - Modest AdaBoost, 105
  - pixel images, 101
- SVHN dataset, 118, 122
- temporal analysis approach, 109
  - Canny rate, 108
  - local appearance, 107–108
  - motion coherency, 107
  - spatial proximity, 106–107
- top-down approach
  - character/noncharacter classifier, 100
  - conditional random field, 98
  - HOGs, 98
  - linguistic context and visual context, 97
  - SWT, 97
  - true detections and false detections, 99
- Torralba's saliency maps, 85
- visual characteristics, 123
- visual saliency models
  - curvature-based approach, 87–90
  - graph-based computations, 84–87
  - probability calculations, 84
- Zhang's fast saliency map, 85
- Script identification
  - candidate text components selection, 212–213
  - Gabor filter, 207, 208
  - gradient histogram method, 210–212
  - intersection points, 213–214
  - language-dependent text detection
    - Bangla and Devanagari text, 197
    - Chinese/English text, 196
    - Farsi/Arabic texts, 196
    - headline-based method, 197–199
    - Indian scripts, 197
    - Korean texts, 196
    - SVM classifier, 196
  - language-independent multi-oriented text detection
    - Bayesian classifier, 200
    - boundary growing (BC) method, 202–203
    - Chinese and Japanese Kanji characters, 200
    - English text lines, 200
    - HHCI, 205, 207
    - HMCI, 205
    - ICDAR-2003 competition data, 205, 209
    - Laplacian/Fourier techniques, 200
    - LMCI, 205, 208
    - LMMI, 205, 206
    - non-straight text lines, 200
    - run lengths, 201–202
    - Sobel/Laplacian methods, 200
    - zero crossing method, 204
  - OCR engine, 207
  - proximity matrices, 214
  - spatial distribution, 213
  - spatial-gradient-features, 210
  - template formation
    - average classification rate, 216
    - confusion matrix, 215, 217
    - sample blocks, 216–217
    - sample templates, 215
- Self-organization map (SOM), 179
- Semantic annotation approach, 228
- Sengamedu, S.H., 190
- Shahab, A., 84, 86
- Shim, J.C., 13
- Shivakumara, P., 14, 30, 31, 52, 58, 88, 93, 96, 200
- Sidhom, S., 228
- Sin, B., 57
- Smith, M.A., 51
- Snoek, C.G.M., 224
- Sobel filters, 23, 24

- Sobel/Laplacian methods, 200
- Sport video analysis systems  
 automatic advertising, 187–188  
 Chinese Professional Baseball League, 185–186  
 content summarization, 187  
 event segment, 184  
 external source information, 183  
 Han's method, 185  
 hidden Markov model, 185  
 image-based features, 185  
 low-level visual/audio cues, 183  
 Major League Baseball, 184  
 mapping, 186  
 object-based features, 185  
 personalized music sports video generation, 187  
 pitch segment, 184  
 scoreboard recognition, 185  
 sport story restoring, 187  
 video content structure, 184  
 webcast text, 185–186
- Sroke width transform (SWT)  
 definitions, 91  
 edge map, 100  
 filling pixels, 94  
 horizontal scene texts, 100  
 true detections and false detections, 99
- Sun, F., 72
- Support vector machines (SVMs), 58
- T**
- Takahashi, M., 185
- Tang, X., 159
- Text-based navigation assistant system, 181
- Text-based person search system, 230
- Text candidates, 129–131
- Text line binarization  
 foreground and background separation, 132  
 low contrast and complex background, 127  
 OCR technology, 127  
 smoothing, 131–132  
 text candidates, 129–131  
 wavelet-gradient-fusion method, 128–129
- Text region extraction methods, 147–148
- Torralba, A., 84
- Torralba's saliency maps, 85
- Tse, J., 147
- U**
- Ultra-mobile personal computer (UMPC), 170–171
- User attention models, 236–238
- US Government Video Analysis and Content Extraction (VACE) program, 249
- V**
- Vasconcelos, N., 240
- Video advertising systems  
 next-generation video ad system, 190  
 online video advertising, 188–189  
 VideoSense, 189–190
- Video character segmentation. *See* Character segmentation
- Video retrieval system, 222
- VideoSense, 189–190
- Video summarization  
 fusing attention models, 236–237  
 fuzzy classification, 236  
 Informedia Project, 236–237  
 multimodal saliency/human annotation, 239  
 POS tags, 238  
 spatial/temporal cues, 240  
 user attention models, 236–238  
 visual attention model, 240
- Video text detection  
 caption text, 3  
 character recognition (*see* Character recognition)  
 character segmentation (*see* Character segmentation)  
 color, 4  
 component-based methods, 6  
 construction and transport industry, 16  
 content-oriented video coding, 15  
 edge and compression, 4  
 feature extraction, 8  
 foreground and background, 7  
 geometry, 4  
 graphics text video fames, 3  
 history  
 automatic sorting machines, 10–11  
 camera-based scene text detection, 13  
 document processing system, 10  
 graphics and scene text, 13  
 graphics documents, 9



- hybrid method, 12
  - image segmentation, 11
  - line drawings, 9
  - OCR, 8–9
  - QBIC system, 11
  - text regularities, 12
  - texture segmentation method, 11
  - top-down vs. bottom-up analysis, 9
  - uniform color, 12
  - internet and digital libraries, 15–16
  - LPR systems (*see* License plate recognition (LPR) systems)
  - machine learning-based methods, 6
  - motion, 4
  - multimedia applications, 8
  - navigation assistant systems (*see* Navigation assistant systems)
  - perceptual video content, 1–2
  - performance evaluation (*see* Performance evaluation)
  - post-processing (*see* Post-processing techniques)
  - preprocessing techniques (*see* Preprocessing techniques)
  - real-time systems, 8
  - scene text, 3–4
  - scene text video frames, 3
  - script identification (*see* Script identification)
  - semantic video content, 2
  - sport video analysis systems (*see* Sport video analysis systems)
  - in supermarkets, 16
  - temporal frames, 6
  - text detection and extraction, 2
  - text localization, 4
  - text racking, 4
  - text recognition, 4
  - text tracking, 4
  - tourist, guide for, 16
  - video advertising systems (*see* Video advertising systems)
  - Video text recognition. *See* Character recognition
  - Visual attention model, 240
  - Visual saliency models
    - curvature-based approach
      - binary and filtered saliency map, 88
      - color saliency, 88
      - curvature saliency, 87
      - gradient vector flow, 88–90
      - Sobel edge map, 88–89
    - text candidates, 89
    - graph-based computations
      - extraction and activation, 84
      - Harel's GBVS model, 86
      - Itti's model, 86
      - Markov process, 85
      - normalization/combination, 84
      - Zhang's model, 87
    - probability calculations, 84
- W**
- Wang, J.R., 187
  - Wang, J.Z., 228
  - Wang, K., 134
  - Wavelet-gradient-fusion (WGF) method, 128–129
  - Wavelet transform techniques, 30–31
  - Webcam, 170–171
  - Web video category/search
    - handheld devices, 242
    - high diversity, 242
    - internet hardware, 243
    - machine learning algorithms, 242
    - mobile devices, 243
    - social information, 242
    - video-sharing websites, 240–241
    - word dictionary, 242
  - Wernicked, A., 249, 251
  - Wolf, C., 159, 248
  - Word and character segmentation
    - binarization method, 148
    - Bresenham's line drawing algorithm, 149
    - character extraction, 153
    - character gap identification, 154
    - error rates, 148
    - Fourier-and moments-based method, 148–149
    - Fourier-moments features, 150–152
    - Monte Carlo sampling, 148
    - multi-oriented text detection, 147
    - scene text detection, 147
    - semantic gap, 147
    - sports events analysis, 147
    - text region extraction methods, 147–148
    - video events analysis, 147
    - word extraction, 150, 152
  - Word extraction, 150, 152
  - Wu, F., 223, 242
  - Wu, L., 106
  - Wu, V., 57, 251
  - Wu, W., 181

**X**

Xiangrong, C., 103  
Xilinx Virtex IV FPGA, 180

**Y**

Yahoo! Video, 188  
Yamamoto, M., 228  
Yan, H., 134  
Yan, J., 134  
Yildirim, G., 100  
Ying, W., 177  
YouTube, 188, 242  
Yu, B., 28, 64  
Yu, D., 134

Yue-ting, Z., 228

Yuille, A.L., 103

**Z**

Zero crossing method, 204  
Zhang, L., 84, 185  
Zhang's fast saliency map, 85  
Zhang, X., 72  
Zhao, W.L., 231  
Zhong, Y., 12, 28  
Zhou, Z., 128, 148, 159  
Zhu, Q., 225, 233, 235  
Zong, Y., 63