学校编码：10384

学号：23220060153376

厦门大学

博士学位论文

基于形状的目标类识别算法研究

The research for shape-based visual recognition of object categories

杜晓凤

指导教师：李翠华

专业名称：控制理论与控制工程

答辩日期：2010年5月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外，该学位论文为(   )课题(组)的研究成果，获得(   )课题(组)经费或实验室的资助，在(   )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。)

声明人(签名):

年   月   日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文(包括纸质版和电子版)，允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

(　　)1.经厦门大学保密委员会审查核定的保密学位论文，于　年　月　日解密，解密后适用上述授权。

(　　)2.不保密，适用上述授权。

(请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。)

声明人(签名)：

年　　月　　日

摘 要

摘 要

视觉目标类识别旨在识别图像中特定的某类目标，基于形状的目标类识别是目前计算机视觉研究的热点之一。真实图像中物体姿态的多样性以及环境的复杂性，给目标的形状提取和识别带来巨大挑战。本文借鉴生物视觉机制的研究成果，对基于形状的目标类识别算法进行研究。主要研究内容如下：

1. 研究与形状认知相关的视觉机制，分析形状知觉整体性的生理基础及其生理模型。以形状知觉整体性为基础，建立基于形状的目标类识别系统框架。框架既重视整体性在自下而上的特征加工中的作用，也重视整体约束在自上而下的识别中的作用。

2. 受生物视觉上的整合野模型启发，本文提出了一个三阶段轮廓检测算法。阶段1利用结构自适应滤波器平滑噪声，增强特征；阶段2通过特征空间上下文的相互作用，调整轮廓响应；阶段3计算轮廓的显著程度。算法重视空间上下文在显著性结构检测中的抑制及增强作用，其中抑制作用实现表面与纹理的分离，增强作用实现同类轮廓特征的结合，从而提高了轮廓检测性能。

3. 研究轮廓的聚合算法。提出基于轮廓生长的连续轮廓提取算法，以显著轮廓片段为种子，兼顾视觉约束与视觉证据，稳定地提取出图像中连续轮廓，减轻了阈值参数对轮廓提取的影响，同时在一定程度上保护低对比度轮廓。改进了利用图像分割提取闭合轮廓的算法，结合形状约束及图像的外观信息进行图像分割，保护形状及外观的连续性，以分割后得到的区域边界作为图像中闭合轮廓。

4. 针对复杂形状的表示，采用分而治之的方法，将轮廓分解成易于表达的轮廓片段。改进形状上下文对单条轮廓片段描述，建立整体与局部的树形关系结构。

5. 在贝叶斯理论框架下，以形状的整体特征为约束，筛选轮廓特征集进行形状匹配。设计了整体优先，部分求精的两级匹配策略。实验表明，引入整体约束的匹配策略，减小了匹配时的搜索空间，提高了匹配精度。

关键词：目标类识别；生物视觉；感受野；轮廓检测；形状匹配

Abstract

Categorical object detection addresses determining the number of instances of a particular object category in an image, and localizing those instances in space and scale. The shape-based visual recognition of object categories is one of hot topics in computer vision. The diversity of poses of targets and complexity of the environment in real images bring huge challenges to shape extraction and object recognition. This work addresses the problems of categorical object detection based on shape by combining ideas and approaches from biological vision. The main parts are as follows.

Introducing the visual mechanisms related to shape perception, interpreting the biological basis for principle of totality in the Gestalt psychology and its biological model. By adopting principle of totality to visual categorical recognition, the bottom-up features extraction and top-down object detection are unified into one framework.

Inspired by integrated field model in biological vision, a three-stage contour detection algorithm is proposed. Stage 1 uses structure adaptive filters to smooth noises and enhance shape features; Stage 2 combines excitatory inputs and inhibitory inputs from spatial context, thus adjusting the contour responses. Stage 3 evaluates the saliency of contour fragments. This algorithm improves detection of object contours and region boundaries in natural scenes by unifying two different roles: excitation and inhibition in the contour detection. Inhibitory interactions are supposed to separate surfaces and textures, while excitatory interactions enhance similar contour feature.

Focused on contour fragments aggregation; a new salient contour extraction algorithm based on adaptive edgel growing is proposed. Starting from seed edgels, construct paths following visual constrains and visual evidences. The progressive growing method makes the algorithm much less sensitive to the threshold parameter and more effective for extraction of low-contrast contours. A method based on image segmentation is developed to extraction closed object contour. Combining the shape cue and appearance cue, segmentation preserves the spatial coherence of the shape and regions appearance. The boundaries of segments are extracted as closed contour finally.

The divide and conquer is used to decompose complex shape into several parts. An extension of the Shape Context descriptor is designed to represent single contour fragment and a tree-union models the relationship between whole and partial feature.

Under the Bayesian theoretical architectures, a coarse-to-fine shape matching strategy is proposed. In both stages, using shape band as global feature constraint, which avoids the disturbances from the other parts on the background and textured edges. Coarse stage evaluates global feature and fine stage determines the contour parts. Experiments show it is simple, yet efficient and effective.

Keywords: Recognition of object categories; Biological vision; Receptive field; Contour detection; Shape matching

参考资料

[1]郑南宁.计算机视觉与模式识别[M]. 国防工业出版社,1998
[2]Schapire R.E., Freund Y., Bartlett P., et al. Boosting the margin: A new explanation for the effectiveness of voting methods [J]. The Annals of Statistics, 26(5):1651-1686, October 1998.
[3]Bayes R. T.. Essay towards solving a problem in the doctrine of chances [J]. Philosophical Transactions of the Royal Society, 1763.
[4]Li. F.F., Fergus R., and Perona P. A Bayesian approach to unsupervised One-Shot learning of Object categories[C]. In Proc. ICCV. 2003
[5]Felzenszwalb P., Huttenlocher D. Pictorial structures for object recognition [J]. Intl. Journal of Computer Vision, 2004, 61(1):55-79.
[6]Crandall D., Felzenszwalb P., Huttenlocher D. Spatial priors for part-based recognition using statistical models[C]. In Proc. CVPR, 2005, 10-17.
[7]Cottes T., Cooper D., Taylor C. et al. Active Shape Models-Their Training and Application[C], In Proc. Computer Vision and Image Understanding, 1995,61(1),38-59.
[8]Fergus R.,Perona P., Zisserman A. A Sparse Object Category Model for Efficient Learning and Exhaustive Recognition [C]. In Proc. CVPR , 2005, 380-387.
[9]Ferrari, V., Fevrier, L., Jurie, F., Schmid, C. Groups of adjacent contour segments for object detection[J].In IEEE Trans. on PAMI (2007).30(1):36-51.
[10]Wu Y., Si Z., Gong H., and Zhu S. Active basis for modeling, learning and recognizing deformable templates [J]. IJCV, 2009.
[11]Shotton J., Blake A., and Cipolla R. Efficiently combining contour and texture cues for object recognition[C]. In Proc. BMVC, 2008.
[12]Opelt A., Pinz, A. and Zisserman A., A boundary-fragment model for object detection[C]. In Proc. ECCV, 2006.
[13]Bai X.,Wang X., Latecki L.J. et.al. Active skeleton for non-rigid object detection[C]. In Proc. CVPR, 2009.
[14]Berg A.C., Berg T.L., Malik J. Shape matching and object recognition using low distortion correspondence [C]. In Proc. CVPR 2005, 1, 26-33.
[15]Zhu Q., Wang L., Wu Y., et al. Contour context selection for object detection: A set-to-set contour matching approach [C]. In Proc. ECCV, 2008.
[16]Marr D., Early processing of visual information [J], Phil. Trans. Roy. Soc. London B, Biol. Sci., 1975,275, 483-524.
[17]Canny J. F., A computational approach to edge detection [J], IEEE Trans. on PAMI, 1986,8(6), 679-698.
[18]Meer P. , Georgescu B., Edge detection with embedded confidence [J], IEEE Trans. on PAMI., 2001.23(12),1351-1365.
[19]Latecki L.J., Sobel M., Lakaemper R., New EM derived from KullbackLeibler divergence[C] , In Proc. ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD), Philadelphia, August 2006.
[20]Fischer B., Buhmann J. M.. Path-Based clustering for grouping of smooth curves and texture segmentation[J], IEEE Trans. on PAMI, 2003 , 25 (4) : 1-6.
[21]Shih F. Y. , Cheng S., Adaptive mathematical morphology for edge linking[C], In Proc. Information Science., 2004. 167,9-21.
[22]Papari G., Campisi P., Petkov N., Multilevel surround inhibition a biologically inspired contour detector[C]. In Proc. SPIE 2007.
[23]Elder J. H., Goldberg R. M., Ecological statistics for the Gestalt laws of perceptual organization of contours[J], Journal of Vision , 2002, 2,324-353.
[24]Papari G., Petkov N., Adaptive Pseudo dilation for gestalt edge grouping and contour detection[J], IEEE Trans. on Image Processing, 17(10), 2008,1950-1962

[25]Kass M, Witkinm A, Terzopoulos D. Snake: Active contour models [J]. IJCV, 1988, 1(4): 321-331.
[26]Neuenschwander, W. Fua, P. Szekely, G. et al. Making snakes converge from minimal initialization[C]. In Proc. ICPR 1994.A:613-615.
[27]Fua P, Brechbuhler C. Imposing hard constraints on soft snakes[C]. In Proc. ICCV, 1996 II:495-506.
[28]Cox, I. J. Rao, S.B. Yu Zhong , " Ratio regions " : a technique for image segmentation[C]. In Proc. CVPR, 1996. 557-564.
[29]Menet S, Saint Marc Phone, Medioni G. B-snakes: Implementation and application to stereo[J]. DARPA image Understanding Workshop, 1990. 720-726.
[30]Flickner M., Sawhney H., Pryor D., and Lotspiech J., Intelligent interactive image outlining using spline snakes[C], In Proc. 28th Asilomar Conf. Signals, Systems, Computers, vol. 1, 1994, 731-735.
[31]Zahn, C.T., Roskies, R.S.: Fourier descriptors for plane closed curves [J]. IEEE Trans. on of Computing. 1972,21(3):269-281.
[32]Gdalyahu, Y., Weinshall, D., Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes [J]. IEEE Trans. on PAMI ,1999,21(12):1312-1328.
[33]Wu J, Leou J. New polygonal approximation schemes for object shape representation [J]. Pattern Recognition, 1993, 26(4):471-484.
[34]Bengtsson A, Eklundh J. Shape representation by multiscale contour approximation[J]. IEEE Trans. on PAMI, 1991, 13(l):85-93.
[35]Hu M. Visual Pattern recognition by moment invariants [J]. RE Trans., Information Theory, 1962, 8(2):179-187.
[36]Shapiro L, MacDonald R, Stemberg S. Ordered structural shape matching with primitive extraction by mathematical morphology [J]. Pattern Recognition,1987,20(l):75-90.
[37]Rosenfeld A. Axial representations of shape [J]. Computer Vision, Graphics, and Image Processing , 1986, 2 (33) : 156-173.
[38]Tari Z S G, Shah J , Pien H. Extraction of shape skeletons from grayscale images [J]. Computer Vision and Image Understanding, 1997, 66 (2) : 133-146.
[39]Belongie S., Malik J., and Puzicha J., Shape matching and object recognition using shape contexts[J], IEEE Trans. on PAMI, 2(4),509-522, Apr. 2002.
[40]Mikolajczyk K., Schmid C., A Performance evaluation of local descriptors [J], IEEE Trans. on PAMI, 2005.27(10), 1615-1630.
[41]Ling H., Jacobs D., Using the inner-distance for classification of articulated shapes [J]. IEEE Trans. on PAMI, 2007. 29(2):286-299.
[42]Felzenszwalb P. F., Schwartz J. D., Hierachical matching of deformable shapes[C]. In Proc. CVPR, 2007.
[43]Ferrari V., Tuytelaars T., and Van Gool L., Object detection by contour segment networks[C], In Proc. ECCV, 2006.
[44]Ferrari V., Jurie F., and Schmid C., From images to shape models for object detection[J]. INRIA Technical Report, July 2008.
[45]Ballard D H. Generalizing the Hough transform to detect arbitrary shapes[J]. Pattern Recognition, 1981, 13 (2) : 111- 122.
[46]Ross M.C., Ronal K., John C. et al. Relation and dynamic time warping: Two methods for tracking ice floes in SAR image[J]. IEEE Trans. on Geosciences and Remote Sensing, 1991, 29(6): 1004-1012.
[47]Thayananthan A., Stenger B., Torr P., et al. Shape contexts and chamfer matching in cluttered scenes. In Proc. CVPR, 2003.
[48]Latecki L.J., Lu C.E., Sobel M., et al. Multiscale random fields with application to contour grouping[C]. In Proc. Neural Information Processing Systems Conf. (NIPS), 2008.
[49]Adluru N. , Latecki L.J. Contour grouping based on contour-skeleton duality[J]. Int. Journal of Computer Vision (IJCV), 2009, 83. 12-29.
[50]Adluru N., Latecki L., Lakaemper R., et al. Contour grouping based on local symmetric[C]. In Proc. ICCV,

2007.
[51]Lu C.E., Longin Jan Latecki, Adluru N., et al. Shape Guided Contour Grouping with Particle Filters[C]. In Proc. ICCV, 2009
[52]Widrow B. The rubber mask technique. Part I and II [J]. Pattern Recognition, 1973, 5:175-211.
[53]Fischler M, Elschlager R. The representation and matching of pictorial structure[J]. IEEE Trans. on Computer, 1973,22(1):67-92.
[54]Bai X., Li Q., Latecki L., Liu W., et al. Shape band: A deformable object detection approach[C]. In Proc. CVPR, 2009.
[55]Palmer S E. Modern Theories of gestalt perception [M]. Oxford, U. K. Blackwell, 1992.
[56]Attneave F. Informational aspects of visual perception [J]. Psychological Review, 1954, 61:183-193.
[57]De Winter, J., Wagemans, J. Segmentation of object outlines into parts: A large-scale, integrative study [J]. Cognition, 2006,99,275-325.
[58]Feldman J., Singh M., Information along contours and object boundaries [J]. Psychological Review 2005, 112(1), 243-252.
[59]Resnikoff H., The illusion of reality: Topics in information science.[M] New York:spinger,1985.
[60]Weisstein N., Harris C.S., Visual detection of line segments: An Object- Superiority Effect [J]. 186(4165), 752-755.
[61]Rosch E. , Mervis C.B. Family resemblances: Studies in the internal structure of categories[J]. Cognitive Psychology, 1975,7:573-605.
[62]http://hi.baidu.com/liujianz/blog/item/653e33b5e72d70ca36d3cab5.html [EB].
[63]Wiesel T N, Hubel D H. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey [J]. J.Neurophysiol,1966,29:1115-1116.
[64]Rodieck RW, Stone J. Analysis of receptive fields of cat retinal ganglion cells [J]. Neurophysiol , 1965(28):833-844.
[65]Lindenbaum M., Fischer M., Bruckstein A.M.. On Gabor ’s contribution to image enhancement [J].Pattern Recognition, 1994,27(11): 1-8.
[66]寿天德,周逸峰. 视觉系统皮层下细胞的方位和方向敏感性[J]. 生物物理学报,1996,48(2): 105-112.
[67]Ghosh, K. Sarkar S.,Bhaumik K.. A possible explanation of the low-level brightness-contrast illusions in the light of an extended classical receptive field model of retinal ganglion cells[J], 2006, 94(2),89-96.
[68]Cai C.F., Liang P.J and Zhang P.M. A Simulation study on the encoding mechanism of retinal ganglion Cell [J]. Lecture Notes in Computer Science, 2007, 4689, 470-479.
[69]邱志诚,黎臧,顾凡及等. 视网膜神经节细胞感受野的一种新模型II[J]. 生物物理学报,2000,16,296-302.
[70]寿天德, 视觉信息处理的脑机制[M],上海科技教育出版社, 1997.12.
[71]潘晴. 生物视觉模型在自动目标识别技术中的应用研究[D].华中科技大学,2007.
[72]Itti L., Gold C., Koch C., Visual attention and target detection in cluttered Natural scenes [J], Optical Engineering, 40(9). 1784-1793, Sep 2001.
[73]Lowe D.G. Object recognition from local scale-invariant features[C]. In Proc. ICCV 1999, 2. 1150-1157.
[74]Kapadia M. K, Westheimer G, Gilbert C D. Spatial distribution of contextual interactions in primary visual cortex and in visual perception[J]. Neurophysiol, 2000, 84 (4), 2048-2062.
[75]Ursino M. , La Cara G.E., A model of contextual interactions and contour detection in primary visual cortex[J], Neural Networks archive, 2004, 17(5-6), 719-735.
[76]La Cara G. E. , Ursino M., A model of contour extraction including multiple scales, flexible inhibition and attention [J], Neural Networks, 2008, 21(5), 759-773.
[77]Grigorescu C., Petkov N. and Westenberg M. A., Contour detection based on non-classical receptive field inhibition [J].IEEE Trans. on Image Processing, 2003,12(7): 729-737.
[78]Grigorescu C., Petkov N. and Westenberg M. A., Contour and boundary detection improved by surround suppression of texture edges[J], Image and Vision Computing, 2004,22 (8), 609-622.

[79]Papari G., Campisi P., Petkov N. et al, A biologically motivated multiresolution approach to contour detection[J], EURASIP Journal on Advances in Signal Processing, 2007, Vol. 2007, Article ID 71828, 28 pages.
[80]唐奇伶,基于初级视皮层感知机制的轮廓与边界检测[D],2007 华中科技大学,博士论文.
[81]Grossberg S., Mingolla E., Neural dynamics of form perception: Boundary completion, illusory figures and neon color spreading [J]. Psychological Review, 1985,92:173-211.
[82]Shen Y. Z., Barner K.E.. Fast adaptive optimization of weighted vector median filters[J]. IEEE Trans. on Signal Processing, 2006, 54(7): 2497-2510.
[83]Donoho D., Johnstone I. Ideal spatial adaptation via wavelet shrinkage [J]. Biometrika, 1994,81(3):425-455.
[84]Ghazel M., Freeman G. H., Vrscay E.K.. Fractal image denoising[J]. IEEE Trans. on Image Processing, 2003, 12(12): 1560-1578.
[85]Tomasi C., Manduchi R., Bilateral filtering for gray and color images[C], in Proc. ICCV 1998, Bombay, India, Jan. 836-846.
[86]Takeda H., Farsiu S., Milanfar P., Kernel regression for image processing and reconstruction applied mathematics and computation [J]. IEEE Trans. on Image Processing, 2007,16(2):419-429.
[87]Zhang J.S., Huang X.F., Zhou C.H.. An improved kernel regression method based on Taylor [J]. Applied Mathematics and Computation. 2008; 193 (2), 419-429.
[88]van Vliet L.J. , FaasF.G.A., Multi-orientation analysis by decomposing the structure tensor and clustering [C],In Proc. ICPR 2006, IEEE Computer Society Press: 856-860.
[89]Harris C. and Stephens M. A combined corner and edge detector [C]. In Proc. Fourth Alvey Vision conference, Manchester, UKm 1988: 147-151.
[90]Kuwahara M., Hachimura K., Eiho S., et al. Digital processing of biomedical images[J]. Plenum Press,New York, NY, 1976. 187-203.
[91]Mahamud S, Williams L R , Thornber K K, et al. Segmentation of multiple salient closed contour from real images[J]. IEEE Trans. on PAMI , 2003, 25 (4) : 433-444.
[92]邹琪,罗四维,钟晶晶.全局显著结构主导下的知觉编组算法[J]. 计算机学报.2007,30(11),2007-2016.
[93]Ren X.F., Fowlkes C. C., Malik J., Learning probabilistic models for Contour completion in natural images [J]. Int. Journal of Computer Vision (IJCV),2008,77: 47-63.
[94]Shin M.C., Goldgof D., Bowyer K.W., An objective comparison methodology of edge detection algorithms for structure from motion task [J], In Empirical Evaluation Techniques in Computer Vision, IEEE CS Press, 1998, 235-254.
[95]Shin M.C., Goldgof D.B., Bowyer K.W., Comparison of edge detectors using an object recognition task [C], In Proc. CVPR, 1999, 1, 360-365.
[96]Shin M.C., Goldgof D.B., Bowyer , Comparison of edge detectors using an object recognition task[C], In Proc. CVIU, 2001 , 84 (1) 160-178.
[97]Grigorescu C., Petkov N., Westenberg M. A., Contour and boundary detection improved by surround suppression of texture edges[J], Image and Vision Computing, 2004, 22 (8), 609-622.
[98]Martin D. R., Fowlkes C. C., Malik J., Learning to detect natural image boundaries using local brightness, color, and texture cues[J], IEEE Trans. on PAMI., 2004, 26(1), 530-549.
[99]Yu S., Shi J. Multiclass spectral clustering[C]. In Proc. ICCV, 2003(1): 313-319
[100]Soundararajan P, Sarkar S. An in-depth study of graph partition measures for perceptual organization. IEEE Trans. on PAMI, 2003 , 25 (6) :642-660.
[101]Ren X.F, Malik J., Learning a classification model for segmentation[C]. in Proc. ICCV 2003, 1, 10-17.
[102]Shi J , Malik J. Normalized cuts and image segmentation[J]. IEEE Trans. on PAMI. 2000,22(8):888-905.
[103]Leung T. , Malik J.. Contour continuity in region-based image segmentation [C]. In Proc. ECCV ,1998, 1, 544-549.
[104]Wu, Q., Yu, Y., Two-level image segmentation based on region and edge integration. In Proc. DICTA , 957-966.
[105]Lakshman P., Alexei N. Skourikhine. Vectorized image segmentation via trixel agglomeration[J]. Pattern

Recognition 2006 39,501-514.
[106]Levinshtein A., Dickinson S., Sminchisescu C. Multiscale Symmetric Part Detection and Grouping[C]. In Proc. ICCV 2009.
[107]Sun Y., Bhanu B., Symmetry Integrated Region-based Image Segmentation [C]. In Proc. CVPR 2009, 826-831.
[108]Gupta L., Mangai U., Das S., Integrating region and edge information for texture segmentation using a modified constraint satisfaction neural network [J]. Image and Vision Computing, 2008,26(8), 1106-1117.
[109]Mignotte M., Segmentation by fusion of histogram-Based k-Means clusters in different color spaces [J]. IEEE Trans. on Image Processing, 2008, 17(5), 780-787.
[110]Hoffman D, Singh M. Salience of visual parts [J]. Cognition,1997,63(l):29-78.
[111]Feldman J., Singh M., Information along contours and object boundaries [J]. Psychological Review 2005, 112(1), 243-252.
[112]Comaniciu D., Meer P., Mean Shift: A robust approach toward Feature space Analysis [J] , IEEE Trans. on PAMI., 2002. 24(5), 603-619.
[113]Ferrari V., Jurie F., Schmid C., Accurate object detection combining recognition and segmentation[C]. In Proc. CVPR, 2007.