

OBJECT CONTOUR COMPLETION BY COMBINING OBJECT RECOGNITION AND LOCAL EDGE CUES

Kar Seng Loke

*School of Information Technology
Swinburne University of Technology Sarawak Campus, Malaysia*

ksloke@swinburne.edu.my

ABSTRACT

We developed a top-down and bottom-up segmentation of objects using shape contours through a two-stage procedure. First, the object was identified using an edge-based contour feature and then the object contour was obtained using a constraint optimization procedure based on the results from the earlier identified contours. The initial object detection provides object category specific information for the contour completion to be effected. We argue that top-down bottom-up interaction architecture has plausible neurological correlates. This method has an advantage in that it does not require learning boundaries with large datasets.

Keywords: Computer vision, object segmentation, object detection, contour extraction, scene interpretation, image understanding.

INTRODUCTION

Detecting and segmenting multiple objects from two-dimensional natural scene images remains a central problem in computer vision as this is often seen as a preliminary step in scene understanding. The problem is compounded by issues of lighting conditions, shadows, occlusion, pose, view angle including accidental matching of the object's surface properties to the background. In recent years, object recognition and localization have achieved very significant results especially with deep neural networks (Olga et al., 2015). However, it has also been shown that deep neural networks are susceptible to

Received: 6 February 2017

Accepted: 28 May 2017

errors that humans would never make (Szegedy et al., 2013, Nguyen & Clune, 2015). This research suggests an alternative for object localization using a more accurate object contour than bounding boxes. Our work also uses shape features instead of the more common image patch features. Image-based cues for recognition have limitations for general categorization as many objects are also identified by their shapes. Shape-based features allow us to recognize objects based on their global features such as shape, and serve the purpose of segmentation as well. Our approach combines top-down and bottom-up cues to extract object contours (rather than bounding boxes) from two-dimensional scene images.

RESEARCH PROBLEM

The aim of this research was to identify the object contours within an image belonging to a natural scene. However, these contours may or may not encapsulate the entire object. Our objective was to extract the object contour as a means of localizing the object.

A basic computational approach is to look for the image edges as discontinuity in intensity. However, an object contour is more than just an intensity differential. Object contours are commonly extracted with an edge detector but the response from such operators is rarely the object's outline contour. This is because an edge in the contour is not only a differential in intensity, but in general, any form of discontinuity between the object and the background. This problem is complicated by camouflage and occlusion. The contour could reflect discontinuities in color and textures as well, and may take into consideration perceptual gestalt properties, prior knowledge and expectation. We wish to be able to extract contour information from a two-dimensional scene image automatically by integrating local and global information sufficiently for image retrieval by object contour. This can be seen as the first step to semantic image retrieval. The difficulty stems from the problem of image understanding in the presence of complex backgrounds.

Identifying object contours can be a chicken and egg problem, since we need to know the object before we can obtain the contours and the contours are used to define the object. Current image contour extraction approaches do not identify the object to extract its contour because the objective is to use those contours to identify the object.

Our main research question centres on precisely how to detect and extract object contours from a natural scene image. We wish to investigate how to create a model for automatic object contour detection and extraction from scene images by using contextual and prior information to improve significantly an object contour-based image retrieval system. We wish to do it with minimal training data. Our approach also takes inspiration from and understanding corresponding neurophysiological and psychological processing in humans and primates.

RELATED WORKS – NEUROPHYSIOLOGICAL RESEARCH

It is known that object recognition is invariant under a variety of conditions, for example, the presence of shadows has little effect on recognition rate (Braje et al., 1998). This suggests a consistent representation that is immune to shadows and argues against an image-based system that performs shadow detection and labeling. As such, local cues would be limited in their use. This suggests that we can rule out using only edges because that would incur the need of differentiating extraneous edges caused by the object, background, texture and shadows. This would also rule out colour, texture and boundary sharpness as sole features used for representations. A preferred approach would then be a global cue that is not easily disrupted locally, e.g. by shadows, such as global shape or object edge contour.

The visual cortex of the brain, located in the occipital lobe (the back of the brain) is responsible for processing visual information. It is known that the response of oriented cells in the V1 and V2 regions in the brain are combined in V4 regions to obtain contour parts information. Cells in the later stages, such as in the IT and lateral occipital complex (LOC) may synthesize V4 signals into global shape and object identity (Grill-Spector & Kushnir, 1998). Human fMRI (functional Magnetic Resonance Imaging) studies suggest that several areas beyond the visual cortex area V4 such as the LOC also participate in object recognition. Various studies indicate that the LOC represents visual objects regardless of inducing cues, size or position. Substantial evidence from fMRI research indicates that the LOC encodes shapes of objects and not low visual features such as textures, contours, color, etc. (Grill-Spector & Sayres, 2008), nor does it encode basic-level semantic categories (Kim et al., 2009).

Various psychophysical researches have shown that top-down facilitation has a role in image segmentation and object recognition. Peterson (1994) suggests that a process called prefigural recognition occurs before figure-ground segregation and before object recognition. In figure-ground analysis, the objective is to separate the object from the background, before the object is recognized. This is generally assumed in the theories of Marr (Marr & Hildreth, 1980). Historically it has been thought that figure-ground categorization precedes object recognition. However, Peterson and colleagues propose that edge detection is performed parallel with figure-ground analysis and simultaneous with the object recognition process.

Vecera and Farah (1997) report that the type and familiarity of the image influence how the image is to be segmented. They also show that by being familiar to the object or shape, the segmentation can override some low level groupings such as connectivity and common region. Their results are consistent with the hypothesis that image segmentation is not a serial process but occurs interactively. They argued that top-down global cues influence bottom-up cues and vice-versa simultaneously and interactively so that object recognition proceeds in pace with figure-ground segregation, not before nor after.

Bar et al. (2006) proposed that the orbitofrontal (OFC) is part of the neurological mechanism that facilitates top-down processing of object recognition. The OFC is activated by a lower resolution image via the magnocellular pathway that connects to the early visual processing areas. The OFC area generates a “coarse” low-resolution interpretation of the input as an initial guess that will be subsequently combined with low-level cues to complete the recognition process. It is suggested that the OFC generates an initial hypothesis based on the coarse aspects of the visual input by serving as a rapid predictor of potential content. Evidences that support the OFC facilitation of early recognition can be found in the review by Fenske et al. (2006). Therefore, it seems that the human visual system co-opts various mechanisms for object recognition, including top-down facilitation.

Psychophysical researches have provided evidence that we are sensitive to curvatures in contours. Experimental work has shown that while performing a visual search task where curved contours are placed amongst straight-line contours, the curved segments pop-out perceptually. The visual system is sensitive to these contours presumably because they are segments of the

contour that contain high amount of information that allow us to quickly recover world structure from the image. De Winter and Wagemans (2008) believed that an important factor for perceptual saliency is the turning angle not the local curvature. The turning angle is measured as the angle between the two flanking lines on both sides of the curve. It is perceptually more salient than its local curvature.

RELATED WORKS – COMPUTATION APPROACHES

The basic component of an object contour is comprised of its edges. A basic computational approach for the image edges is discontinuity in intensity. However, an object contour is more than just an intensity differential. Object contours are commonly extracted with one of the edge detector processing algorithms such as the Canny edge detector (Canny, 1986). The response from such operators is rarely the object's outline contour. This is because an edge in the contour is not only a differential in intensity, but, in general, also any form of discontinuity between the object and the background. This problem is complicated by camouflage and occlusion. The contour could reflect discontinuities in color and textures as well, and may take into consideration perceptual gestalt properties, prior knowledge and expectation. Contour detection is a global concept related to the meaning and recognition of the object from the ground. The challenge is to integrate all these considerations into a viable model.

We briefly summarize and group the common approaches to object segmentation into three categories: edge-based (Marr & Hildreth, 1980), contour-based and region-based. Edge-based approaches, as the name implies, use edge detection and edge linking to segment an object. The linking of edges is usually based on psychological and gestalt properties (Shashua & Ullman, 1988). A contour-based approach would frequently include an active contour paradigm (Caselles et al., 1997). Region-based approaches primarily include region clustering, graphical methods, diffusion (Perona & Malik, 1990) and variational (Shi & Malik, 2000) methods. Texture information may also be used to determine legitimate object boundaries and edges (Chaji & Ghassemian, 2006). Another approach takes inspiration from biology, mimicking the center-surround receptive field neurons of the human visual system (Papari et al., 2007). These edge detectors provide a set of primitives that can be used to form more elaborate models using chains, lines, circles and splines.

Image segmentation can also be posed as a graphical Bayesian problem. Xiaofeng Ren et al. (2005) used constrained Delaunay triangulation to enforce curvilinear continuity with loopy belief propagation to drive edge segmentation. Felzenszwalb et al. (2006) used the Markov process to find salient curves and suppress noisy edges. Recently structured learning that incorporates random forests has been used to predict local edges (Dollar & Zitnick, 2013).

There are several approaches that use visual contours to classify objects. The Shape Band approach (Bai, Li et al., 2009) uses a coarse-to-fine approach to determine the object contour. The Shape Band defines a radius distance from the image sampled edge points from which approximate directional matching of points could be performed. Edges within the Shape Band would be then matched more accurately using Shape Context (Belongie, Malik et al. 2002). Ferrari's (Ferrari, Tuytelaars et al., 2006; Ferrari, Jurie et al., 2010) work used a local feature which they called pairs of adjacent segments (PAS). Each pair of connected segments forms one feature set that includes the mean of the two segment centres, distance between the segment centres, edge strength, a descriptor that encodes the shape of the PAS using the segments' orientations and lengths, and the relative location vector. A codebook is created by clustering the PAS inside all training bounding-boxes according to their descriptors. These will be used for matching or recognizing object shapes.

Other schemes used a dictionary of contour fragments (Opelt et al., 2006; Arandjelovic & Zisserman, 2011; Haribaran et al., 2011) that is learnt for classification. Ben-Yosef (Ben-Yosef et al., 2015) used local features based on points, contours and regions (e.g. ear, neck) and their spatial relations to identify local semantic elements, which can then be combined to produce a global interpretation.

METHODOLOGY

In Loke (Loke et al., 2010), an approach that uses a two-stage procedure to segment an image following prior recognition has been proposed. The approach taken is to acquire a preliminary identification of the object initially, followed by segmentation based this prior knowledge (Figure 1). Instead of image patches normally used in object recognition, edge contours are used as initial features for pre-identification and they are used as a guide for segmentation.

The object recognition phase works by detecting edges using the standard intensity-based edge detection technique (i.e. Canny edge detection). From the edges detected, we discarded the shorter edges. The longer edges that are kept are analyzed for turning points.

We used turning points as representation of contour fragments (Loke, 2013). The contour fragments were extracted using a standard edge extraction algorithm from which the turning points were extracted. These extracted turning points (Figure 2) were compared with a collection of compiled exemplar image turning points. In the recognition process, scale and position were accounted for using a variable size sliding window approach. Windows of various sizes were slid across the target scene image, at each window location; the number of matching turning points was recorded. A bounding box was drawn over the window where the probable location of the object was. This was determined by the window that had the greatest matching turning points.

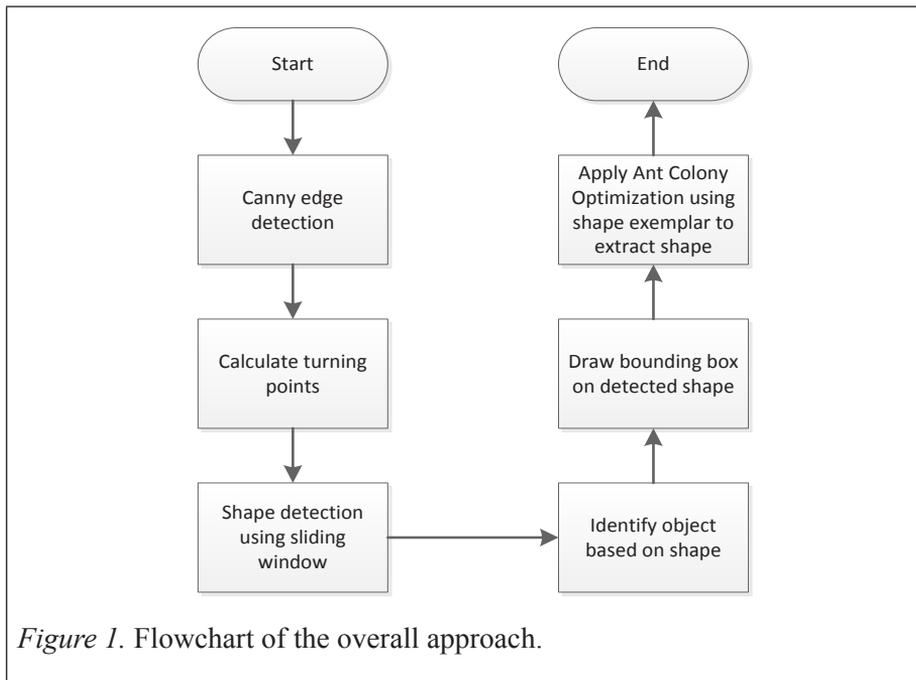


Figure 1. Flowchart of the overall approach.

When the object had been preliminarily recognized within the bounding box (Figure 3), we would have had a set of intensity edges within the window. Some of these edges had been identified to be part of the object contour when performing the recognition process. However, these edges are not definitive,

as some of them may have been misconstrued due to process errors, object overlap or background confusion and so on. We knew that likely that a majority of the edges belonged to the object or else we would not have been able to identify them earlier. Having identified them we would know the probable shape of the object. With this preliminary information and constraint, we wanted to find a contour that fitted the image given. This could be seen as a constraint satisfaction problem but with the constraints themselves as being uncertain. We wanted to optimize the contour path fitness against the exemplar versus the number of edges in the image. As the first approach, we used the ant colony optimization (ACO) algorithm (Dorigo & Stützle, 2004) to investigate this problem.

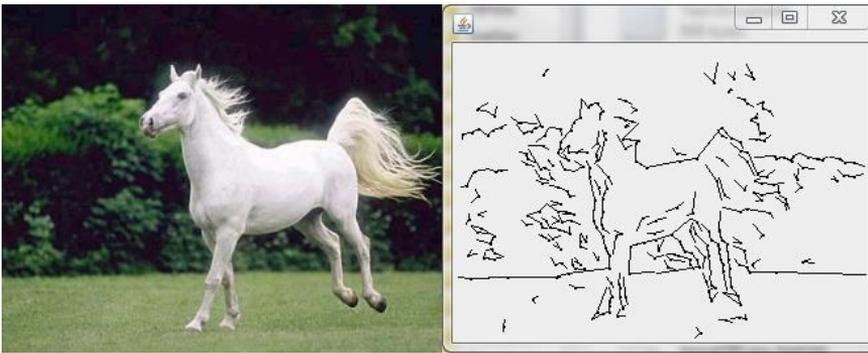


Figure 2. (Left) Original image. (Right) The detected turning points (point & 2 lines) of the image.



Figure 3. Bounding box identification of horse based on shape using turning point contour fragment.

IMPLEMENTATION

The image was divided into 8x8 blocks. Each block had a corresponding node in the ACO network. Each ACO node was seeded according to the edge strength in the corresponding image block. The edges not marked in the recognition process had their initial pheromone strength reduced by percentage that was determined empirically.

The ants in the ACO algorithm needed to discover and build a path that corresponded to the object contour. Apart from the constraints given in the ACO algorithm, we wanted to impose a few general constraints when constructing the path. The reason was to reduce the degrees of freedom available to avoid creating undesirable paths. The constraints we used were:

1. The path should be as smooth as possible. The reason this was desired was to prevent convoluted paths that fold back and forth, and natural objects generally have smooth curves.
2. The path should not crossover or turn back on itself. This was to prevent creating a space-filling path with no internal spaces.
3. The start and end node should be close to each other. This was ensure a closed loop contour.

To satisfy constraint (1) we simply required that the constructed path follow in the same general direction as much as possible. For example, if the path was previously heading from right to left, then it should be constrained to only three possible choices on the left (the shaded squares in Figure 4) out of the 8 possible neighbouring points.

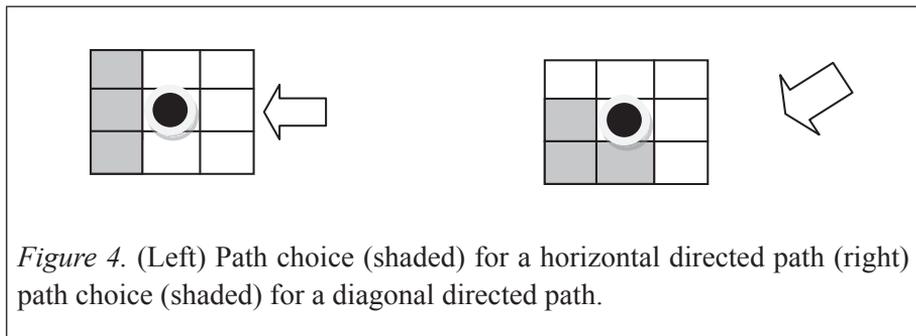


Figure 4. (Left) Path choice (shaded) for a horizontal directed path (right) path choice (shaded) for a diagonal directed path.

When the directed choice path points were blocked, where the path had reached the border, or crossed back to itself, then only other path choices

would be made available. For instance, in the above example, if the shaded points were not available, then the non-shaded points would be considered.

We also needed to check that the path did not cross back on itself, and that it did not get too close to itself. This was accomplished by checking that the possible neighbouring points distance exceeded a certain distance threshold from all its previous path points. Points that were too close to existing path points were not permitted.

These constrained available choices were then passed on to the ACO algorithm to pursue based on its optimization scheme of exploitation or biased exploration. The desirability of a path point (or node) was based on the original edge strength and deposited pheromones. Finally, the path was terminated if the next point was close to the starting point.

After the path was constructed, the quality of path was evaluated. We wanted to evaluate the quality of the path by its similarity to the shape of the exemplar, which was the object that had been pre-identified via recognition. The exemplar was the model shape that would be used to guide the object segmentation. Basically the path should be converging to the exemplar while being constrained by the actual edges. This implies that the path quality measure should be some shape fidelity measure. At the same time, the path quality also includes a measure of how many detected edges in the recognition are counted as part of the path.

For this measure, we performed chamfer matching (Thayananthan et al., 2003) using the distance transform. Chamfer matching is a popular technique used to find alignment between two edge maps. In our case, we used one edge map from the exemplar (Figure 5), and used that to match the ant-constructed path. Chamfer matching provides a continuous measure that can tolerate small misalignments, occlusions and deformations. The chamfer distance is efficiently computed via the distance transform. The distance transforms (DT) calculate each point x , the distance to the nearest edge, x_E :

$$DT_E(x) = \min_{x_E \in E} \|x - x_E\| \quad (1)$$

The chamfer distance gives the average distance from the points, x_p in the ant-constructed path to the closest exemplar edge. The chamfer distance is then simply given by:

$$D_{chamfer}(x) = \frac{1}{|x_p|} \sum_{x_p} DT_E(x_p) \quad (2)$$

This is simply a lookup operation given the pre-calculated distance transform from the exemplar. This distance has to be calculated each time an ant constructs a path, so it needs to be efficient. However, we need to compare it in a scale-invariant manner, since the target object can be of different sizes. Therefore, before calculating the chamfer distance, we rescaled the size of the path to the size of the exemplar template (Figure 5).

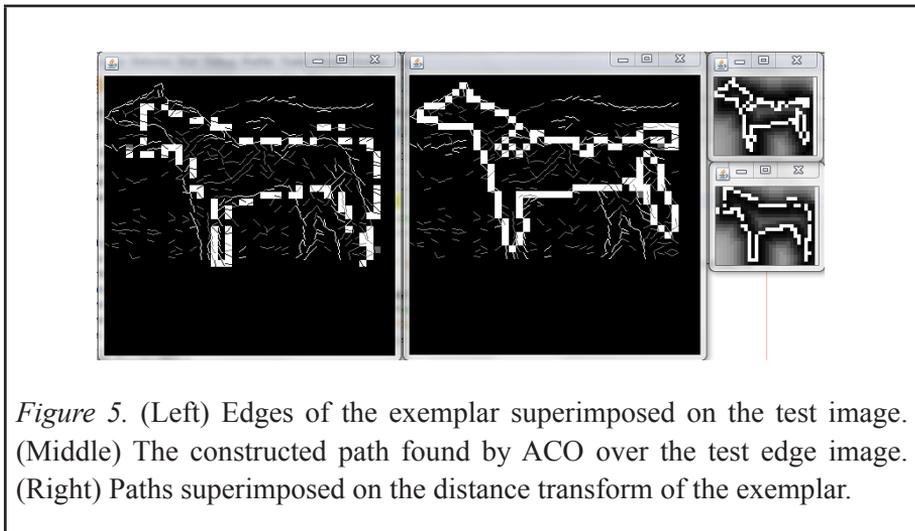
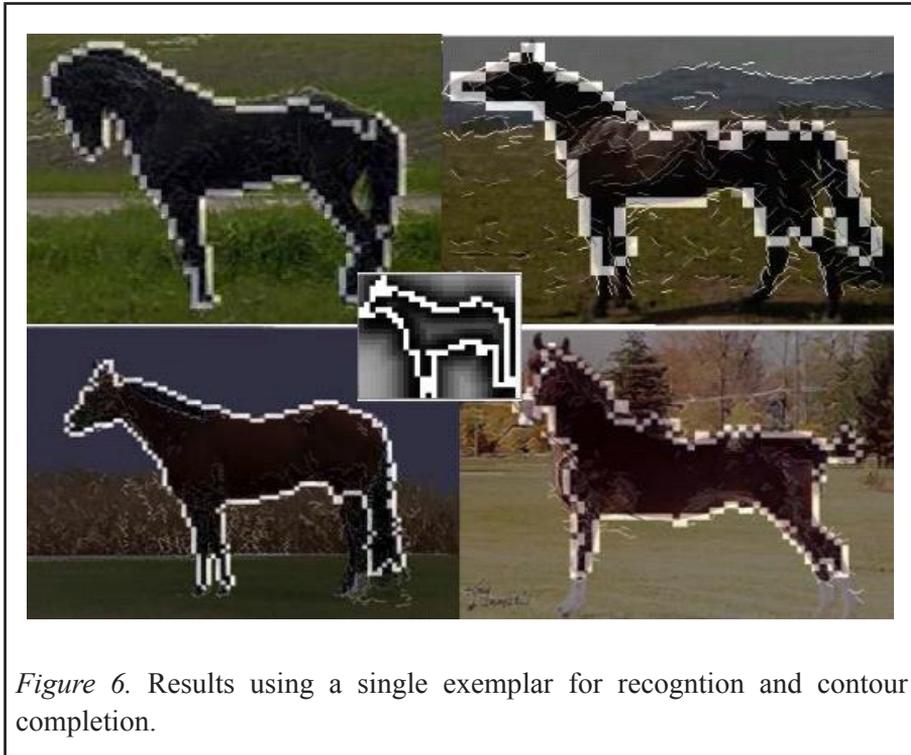


Figure 5. (Left) Edges of the exemplar superimposed on the test image. (Middle) The constructed path found by ACO over the test edge image. (Right) Paths superimposed on the distance transform of the exemplar.

For the ACO algorithm, we used 20-40 iterations per epoch, of 1500-3000 epochs. At the end of each epoch, the best ant is allowed to update the pheromones on its path. If this is the best all time ant, then it updates the full amount, else a fraction of the full amount is applied (we used 0.5-0.7 of the full amount). The update amount is the scaled value from the chamfer distance value. At each epoch, all nodes in the ant network are evaporated.

RESULTS

Figure 6 shows samples of the ant constructed path generated from a single exemplar shape. This shows that the algorithm can generate flexible solutions according to the constraints and varying shapes.



The results in Figure 6 show that it is possible to obtain coarse contours in a large scale invariant manner (Figure 6; all images are of different sizes) despite using only one exemplar. The results show clearly the shape of a side-facing horse. Though the results are not perfect, it is possible with some additional post-processing, such as smoothing and constructing a convex path, some defects can be ameliorated. A typical number of paths constructed and examined was in the order of about 10,000 paths.

There are some weaknesses in the ACO approach. For example, for larger images, the search space is too large for the ACO, and it becomes sensitive to initial conditions. The path contour quality evaluation is critical in determining its success. The chamfer distance was calculated using the average – this meant that some paths could give the same results even though the paths differ, this is as long as the further path points get cancelled out, returning the same average. The other error was due to the rescaling error. The exemplar that we used was smaller than the image being tested.

The results could be improved with better path and shape measures. However, we did not have resources to test other evaluation functions. The results show that it was possible to recover contour paths, and that was sufficient to make our case. To deal with shapes with greater variability such as different poses, viewpoints or with moving appendages, simultaneous multiple exemplars could be used (Figure 7).

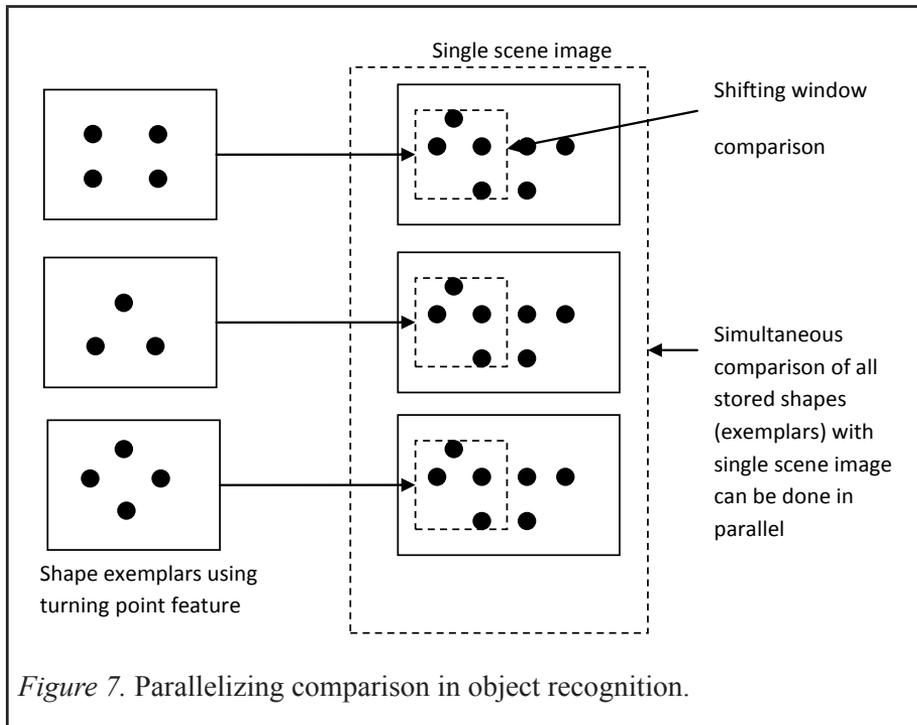


Figure 7. Parallelizing comparison in object recognition.

Our approach (Figure 8), in general, worked this way: from a particular to a scene image, we attempted to determine what object was in the image, and having determined the object, we would use the object shape to guide the contour extraction. The next stage would include the feedback stage where the quality of the contour was determined. If the contour fitted the constraints well, then the task was complete or else another shape exemplar would be used. We did not implement the feedback model, but we demonstrated the top-down bottom-up contour extraction.

The comparison stage used a shifting window to compare the salient turning points from a set of reference image turning points. This process is fast

and can be parallelized because each shape and window can be compared independently.

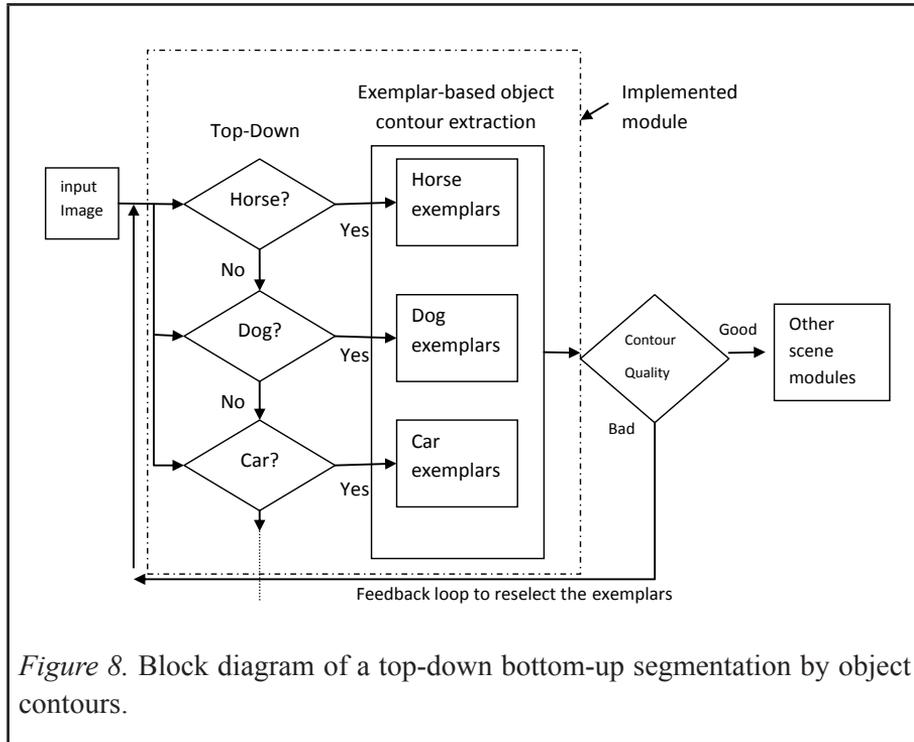


Figure 8. Block diagram of a top-down bottom-up segmentation by object contours.

CONCLUSION

We implemented the extraction of contours that can be facilitated by top-down knowledge in a computational program. We have shown that this approach is viable by proof of concept. However, our approach is not meant to be biologically accurate but to be inspired by it. This approach has the advantage that contour construction is based on a single per pose exemplar identified by an earlier recognition stage. It does not require learning boundaries from many training samples. The learning is only done in the earlier recognition stage. Putting all these together, we proposed a model that has a recognition phase that precedes actual object contour segmentation. This model is based on experimental and theoretical considerations. Psychological and neurophysiological research suggests that recognition may occur interactively guided by top-down processes from bottom-up cues. We discussed the

literature that demonstrated turning points as biologically salient points. We used these biologically salient turning points for preliminary scale-invariant object recognition or detection.

We would work on formalizing this approach and search for better shape measures. A more formal approach will better facilitate analysis. Object detection and recognition have been massively successful using deep learning approaches (LeCun et al., 2015). Incorporating deep neural networks would also be an approach we would take in the future.

REFERENCES

- Arandjelovic, R., & Zisserman, A. (2011). Smooth object retrieval using a bag of boundaries. *Proceedings of the 2011 International Conference on Computer Vision*, 375-382.
- Bai, X., Li, Q., Latecki, L.J., Liu W., & Tu, Z. (2009). Shape band: A deformable object detection approach. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2009.5206543>
- Bar, M., & Kassma, K. S. et al. (2006). Top-down facilitation of visual recognition. *Proc. Natl. Acad. Sci USA*, 103(2), 449-454. doi: 10.1073/pnas.0507062103
- Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 24, 509-522. doi: 10.1109/34.993558.
- Ben-Yosef, G., Assif, L., Harari, D., & Ullman, S. (2015). A model for full local image interpretation. In Noelle, D. C., Dale, R., Warlaumont, A. S., Yoshimi, J., Matlock, T., Jennings, C. D., & Maglio, P. P. (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 220-225.
- Braje, W. L., Kersten, D. et al. (1998). Illumination effects in face recognition. *Psychobiology*, 26(4), 371-380. doi:10.3758/BF03330623

- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 8(6), 679-698.
- Caselles, V., Kimmel, R. et al. (1997). Geodesic active contours. *International Journal in Computer Vision*, 22(1), 61-79.
- Chaji, N., & Ghassemian, H. (2006). Texture-gradient-based contour detection. *The European Association for Signal Processing (EURASIP). Journal on Applied Signal Processing* 21709, 1-8. doi:10.1155/ASP/2006/21709
- Dollar, P., & Zitnick, C.L. (2013). Structured forests for fast edge detection. *Proceedings IEEE International Conference on Computer Vision*, 1841-1848. doi: 10.1109/ICCV.2013.231
- Dorigo, M., & Stützle, T. (2004). *Ant colony optimization*. MIT Press.
- Felzenszwalb, P., & McAllester, D. (2006). A min-cover approach for finding salient curves. *IEEE Computer Society Workshop on Perceptual Organization in Computer Vision*, 185-185. doi: 10.1109/CVPRW.2006.18
- Fenske, M. J., Aminoff, E. et al. (2006). Top-down facilitation of visual object recognition: Object-based and context-based contributions. *Progress in Brain Research*, 155, 3-21. doi: 10.1016/S0079-6123(06)55001-0
- Ferrari, V., Jurie, F. et al. (2007). Accurate object detection with deformable shape models learnt from images. *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition*. doi: 10.1109/CVPR.2007.383043
- Ferrari, V., Jurie, F. et al. (2010). From images to shape models for object detection. *International Journal in Computer Vision*, 87, 284. doi:10.1007/s11263-009-0270-9
- Grill-Spector, K., & Sayres, R. (2008). Object recognition - Insights from advances in fMRI methods. *Current Directions in Psychological Science*, 17(2), 73-79. <https://doi.org/10.1111/j.1467-8721.2008.00552.x>

- Grill-Spector, K., Kushnir, T. et al. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping, 6*, 316-328.
- Hariharan, B., Arbeláez, P., Bourdev, L., Maji, S., & Malik, J. (2011). Semantic contours from inverse detectors. *Proceedings IEEE International Conference on Computer Vision*, 991-998. doi:10.1109/ICCV.2011.6126343
- Kim, J. G., Biederman, I. et al. (2009). Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics. *Vision Research, 49*, 2297-2305.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.
- Loke K.S., & Egerton, S. (2010). Scene understanding: A framework for image segmentation via object recognition. *Proceedings of the Sixth International Conference on Intelligent Environments*, 328-331. doi: 10.1109/IE.2010.67
- Loke, K. (2013). Contour-based shape recognition using perceptual turning points. *Proceedings of the International Conference on Computer Vision Theory and Applications - Volume 1: VISAPP, (VISIGRAPP 2013)* ISBN 978-989-8565-47-1, pages 487-491. doi: 10.5220/0004304804870491
- Marr, D., & E. C. Hildreth (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences, 207*(1167), 187-217.
- Nguyen A, Yosinski J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. *Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR '15)*, 427-436.
- Olga Russakovsky, Jia Deng, Hao Su, Jonathan, K. Sanjeev S. Sean Ma, Zhiheng Huang, Andrej, K. Aditya, K. Michael, B. Alexander, C. B. & Li Fei-Fei. (2005). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision, 115*(3), 211-252.

- Opelt, A., Pinz, A., & Zisserman, A. (2006). A boundary fragment model for object detection. In Leonardis, Aleš, Bischof, Horst, Pinz, Axel (Eds.), *Computer Vision -- ECCV 2006, 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, Proceedings, Part II*, 575-588. doi: 10.1007/11744047_44
- Papari, G., Campisi, P., Petkov, N., & Neri, A. (2007). A biologically motivated multiresolution approach to contour detection. *EURASIP Journal on Advances in Signal Processing*, 1-28. doi: 10.1155/2007/71828
- Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 12(7), 629-639.
- Peterson, M. A. (1994). Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science*, 3, 105-111.
- Ren, X., Fowlkes, C., & Malik, J. (2005). Scale-invariant contour completion using conditional random fields. *Tenth IEEE International Conference on Computer Vision (ICCV'05)*, 2, 1214-1221. doi: 10.1109/ICCV.2005.213
- Shashua, A., & Ullman, S. (1988). Structural saliency: The detection of globally salient structures using a locally connected network. *Second International Conference on Computer Vision*, 321-327. doi: 10.1109/CCV.1988.590008
- Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 22(8), 888-905.
- Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). *Intriguing properties of neural networks*. Retrieved from: [http:// arXiv preprint arXiv:1312.6199](http://arXiv.org/abs/1312.6199).
- Thayananthan, A., Stenger, B., Torr, P.H.S., & Cipolla, R. (2003). Shape context and chamfer matching in cluttered scenes. *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*, 1, 127-133. doi: 10.1109/CVPR.2003.1211346

Vecera, S. P., & Farah, M. J. (1997). Is visual image segmentation a bottom-up or an interactive process? *Perception & Psychophysics, 59*(8), 1280-1296.

Winter, J. D., & Wagemans, J. (2008). Perceptual saliency of points along the contour of everyday objects: A large-scale study. *Perception & Psychophysics, 1*(70), 50-64.