Detection of Multi-part Objects by Top-down Perceptual Grouping

Vénérée Randrianarisoa, Jean-François Bernier, and Robert Bergevin

Laboratoire de Vision et Systèmes Numériques (LVSN), Pavillon Adrien-Pouliot
Université Laval, Sainte-Foy (Qc), Canada, G1K 7P4
E-mail: veneree@gel.ulaval.ca, bergevin@gel.ulaval.ca

Abstract

In this paper, a top-down approach based on perceptual grouping is proposed for multi-part objects detection. The abstract conceptual category of multi-part objects is formalized by a set of global criteria. These criteria will enable the evaluation of the segmentation quality in order to determine if the whole grouping is perceptually significant and if it has a good perceptual shape. A new cognitive vision methodology, called SAFE (Subjectivity And Formalism Explicitly), is presented. Its goal is to help identify the proper global criteria and to validate the judgment derived from formal calculations of these criteria by human judgment.

Keywords: perceptual grouping, top-down processing, methodology, object detection.

1. Introduction

The goal of image segmentation is to partition the image in order to identify structures that correspond to meaningful objects. One of image segmentation techniques is based on perceptual grouping, which can be defined as the ability of the visual system to organize primary data as pixels, segments or regions present in the observed scene into perceptually significant groupings. In the following, we will refer to these primary data as primitives.

Perceptual grouping relies on a set of properties, known as Gestalt laws or perceptual grouping cues, the most common being parallelism, symmetry, closure, proximity, similarity, smoothness [7]. Lowe [9] introduced the use of grouping for object recognition in the 1980's.

Though great works have been done, our review of perceptual grouping literature has shown that only a few ideas of the Gestalt movement have been exploited. Few works incorporate the principle of Prägnanz, in spite of its importance. Prägnanz was defined as the tendency of humans to group primitives in accordance to criteria such as regularity, smoothness and visual balance. On the other hand, the Gestalt basic principle is that the whole (e.g. an object like a car) carries a greater cognitive significance than its individual components. In a way, these two principles permit to give meaningful global quality to the whole grouping. Nevertheless, most of the proposed methods in perceptual grouping do not consider this global concept because they follow a bottom-up process.

Indeed, most works of perceptual grouping are based on a bottom-up process [12]. In this approach, the primitives are grouped in a hierarchical way according to the Gestalt laws without a priori knowledge of the image contents [13], [4], [11].

Top-down segmentation looks for a predefined object in the input image and is guided by a priori knowledge about the object, such as its shape, color, or texture. Research works in this area are essentially model-based. In the domain of perceptual grouping, past efforts have been made to form simple and small groupings such as convex groupings, rectangles or simple polygons. They are then aggregated in order to extract large and complex groupings that should correspond to meaningful objects in the image [11]. Jain *et al.* [6] observed the lack of top-down processing that incorporates Gestalt principles for object recognition, in the state-of-the-art of computer vision. This is due, partly, to the fact that top-down process to obtain directly large and complex groupings has a very high computational cost. For instance, the dimension of the search space to find such groupings in a set of primitives grows exponentially with respect to the number of primitives.

However, in our application, it is needed to extract such large and complex groupings. Our aim is to design an algorithm for the detection of multi-part objects. In this paper, we present a top-down approach based on Gestalt principles to extract their boundary.

Related works on object boundary extraction based on perceptual grouping exist in the literature [14], [5], [3], [2], [10]. They use essentially Gestalt laws of closure, continuity, proximity and convexity. The objects of interest are simpler than ours, since these algorithms work well for natural, smooth and compact objects. Moreover, the grouping cues are defined intuitively and are never validated by human observation.

The main contributions of this paper include: (i) a top-down approach that relies on the identification of global criteria for an abstract object category, more specifically multi-part object category, by subjective human observations, (ii) an incorporation of the principles of Prägnanz and Gestalt laws into these global criteria, (iii) a formalization of these global criteria, (iv) a new methodology, referred to as SAFE (Subjectivity And Formalism Explicitly), that helps identify and validate the global criteria.

This paper is organized as follows. In Section 2, we describe the SAFE methodology and in Section 3, we present its application to the detection of multi-part objects.

2. The SAFE methodology

A well-known methodology for model-based object recognition utilizes a supervised learning from a large set of sample objects. A pattern representing a set of features is used and decision-theoretic methods are applied according to the statistical results of the learning.

The methodology presented in this paper is different since it is based on an a priori identification of global criteria corresponding to a given abstract conceptual category. We believe that these global criteria should be similar to those used by human visual system. Indeed, although sophisticated artificial visual systems exist, no one can perform as well as humans to detect objects. In our approach, we have to emphasize that such human capabilities are integrated in the design of our algorithms, which is fundamentally different to existing a posteriori psycho-evaluations of segmented images.

A flowchart of the methodology is presented in Fig. 1. For our application, the input data are a set of N constant curvature primitives (CCPs), $P = \{p_j, j = 1..N\}$, where p_j is a primitive (arc or line segment), extracted from an image that may contain one or many multi-part objects.

This methodology is applied to top-down segmentation under the following hypotheses:

- nature of objects: multi-part objects
- low-textured objects
- relatively low-cluttered background

In order to apply the methodology, we have designed an interactive software tool, that we call SAFE-T for SAFE-Tool, with a graphical user interface (see Fig. 2).

The methodology consists of four steps (SGT, FGT, FV, SV): its goal is to validate the judgment derived from formal calculations which we refer to as FGT (Formal Ground Truth) by human judgment which we refer to as SGT (Subjective Ground Truth).

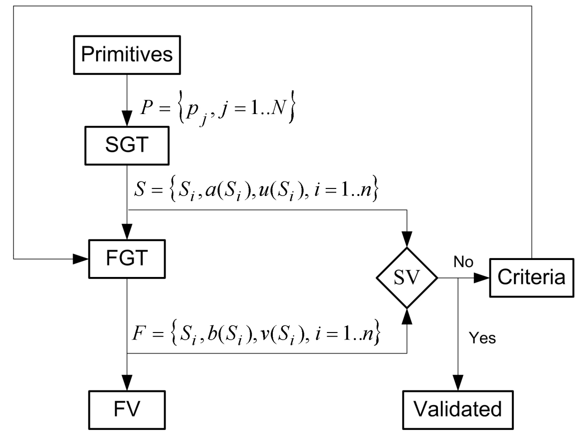


Figure 1. The SAFE methodology.

Subjective Ground Truth (SGT): In this first step, a list of groupings or subsets from P is generated. This list corresponds to possible solutions of the image segmentation and may reflect a certain degree of perturbation around the best solution (see Fig. 2).

A subset is formed as an ordered sequence of primitives, e.g., selected manually by means of SAFE-T. A subset $S_i \subset P$ is defined as $S_i = \{p_j, j = 1..n_i\}$, where the number of primitives n_i can be different from a subset to another ($n_i \geq 2$).

Then, the operator is asked to evaluate the perceptual quality of each subset of the list by answering the question: "How good is this subset as a multi-part object boundary ?". For each subset S_i, he is asked to give a score, referred to as subjective score and noted $u(S_i)$, in accordance to his own subjective judgment of the subset. The subset corresponding score represents the relative quality of the solution.

Once subjectively evaluated, the list of subsets is sorted by decreasing value of $u(S_i)$. Therefore, the best evaluated

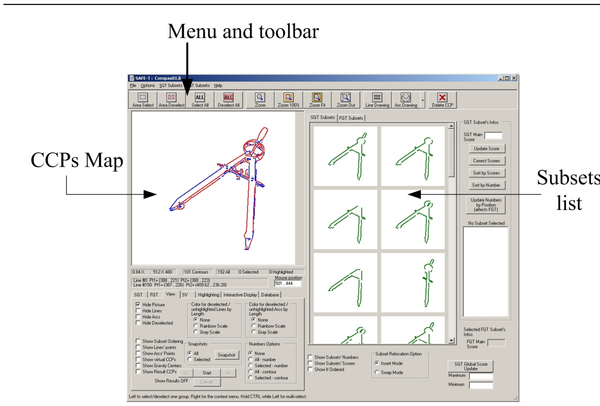


Figure 2. SAFE-Tool graphical user interface.

solution is ranked first. The SGT step will yield a sorted subjective set, $S = \{S_i, a(S_i), u(S_i)\}$, where $a(S_i)$ is the subjective rank of the subset (S_i). A subset which has been judged according to the subjective score $u(S_i)$ will be called subjective subset.

At the end of this step, S should reflect the perceptual properties of an hypothesized multi-part object. A verbal description of the abstract conceptual category of the object is then identified, leading to a set of global criteria to be optimized for the multi-part object detection. This step is repeated for several images with different multi-part objects in order to refine and complete the set of criteria.

Formal Ground Truth (FGT): In this second step, formal model of each global criterion is proposed. These criteria are combined in a unique measure, noted $v(S_i)$, and referred to as formal main score in the rest of the paper. The list of subsets defined in SGT step is evaluated objectively by means of the formal main score. This list is then sorted by decreasing value of $v(S_i)$ and a sorted formal set, noted $F = \{S_i, b(S_i), v(S_i)\}$ is obtained, where $b(S_i)$ is the formal rank. A subset which has been judged according to the formal score $v(S_i)$ will be referred to as formal subset.

Formal validation (FV): This step consists in verifying the implementation of each global criterion. The validation could be applied on either synthetic or real data.

Subjective validation (SV): The aim of this step is to validate the formal model of the global criteria by human judgment. This is achieved by comparing the two rankings defined in SGT and FGT. The criteria formalization and implementation are validated when the list of subsets has the same ranking in SGT and FGT, i.e. $b(S_i) = a(S_i)$. If this is the case, the segmentation algorithm design is achieved, otherwise the formalization or the implementation of the algorithm must be corrected. This leads to a closed-loop process.

3. Application

By following the SAFE methodology, we have identified a set of global criteria that correspond to the formalization of the abstract concept category of multi-part objects. We have observed that these criteria are in accordance to the theory of Prägnanz and to Gestalt laws.

3.1. Criteria formalization

The criteria values are normalized so as to be in the range [0,1].

As some criteria need to be computed on closed boundaries, we fill the gap between two oriented primitives by joining their two endpoints with a virtual line.

3.1.1. Number of parts

This criterion is important to our application since we are looking for multi-part objects. In this paper, we do not want to determine a precise decomposition of an object into its constituent parts because we are interested in a general form of a given object category. Therefore, an approximate number of parts is sufficient.

In order to determine approximately the number of parts of the object, we propose to use concavities on the object boundary, which are present at significant negative curvature points. These dominant points are interesting, as they are situated at significant visual part cuts. However, their saliency is not sufficient to obtain part decomposition since they may correspond to noise. In order to overcome this problem, we propose to use the relevance measure proposed in [8], and apply two kinds of recursive filtering based on this measure so that significant noises on the boundary are eliminated without modifying its global shape.

This measure, noted $K(p_1, p_2)$ depends on the turn angle and the length of two consecutive primitives (p_1, p_2) and is computed as:

$$K(p_1, p_2) = \frac{\beta(p_1, p_2) l(p_1) l(p_2)}{l(p_1) + l(p_2)} \quad (1)$$

where l is the length of the segment and $\beta()$ is the turn angle between (p_1, p_2).

The definition of the turn angle is shown in Fig. 3), where $T_2(B)$ is the tangent direction that opens the second arc and $T_1(B)$ the one that closes the first arc, both at common point B.

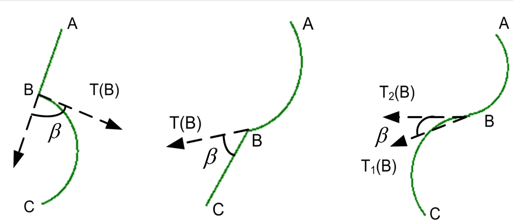


Figure 3. Turn angles for arc and line segment.

The dominant points are determined as those situated at points where the saliency of two consecutive primitives computed by $K(p_1, p_2)$ is negatively high. The first filtering is to eliminate irrelevant turn angle, i.e. very small value of turn angle: this is done by evolving the boundary until the minimum value of turn angle is superior to a predefined threshold Θ_1.

Furthermore, we observe that, for the same value of $\beta(p_1, p_2)$: if $l(p_1) \gg l(p_2)$, then $K(p_1, p_2) \approx \beta(p_1, p_2)l(p_2)$; and if $l(p_1) \approx l(p_2)$ then $K(p_1, p_2) \approx \frac{\beta(p_1,p_2)l(p_2)}{2}$.

It means that, for the same value of turn angle, even when the length of one of the primitive is very large, $K()$ is only influenced by the length of the shortest one. Furthermore, if we compare the two cases, a ratio of 2 always exists between the two values of $K()$ whatever the length of the longest one, i.e. they are not normalized to the total length of the two primitives.

The second filtering takes into account this length influence and linearizes the boundary if one of the following cases occurs: (i) $l(p_1) \leq L$ and $l(p_2) \leq L$ (ii) $l(p_1) \leq L$ and $l(p_2) \geq R \cdot L$ and $\beta(p_1, p_2) \leq \Theta_2$ (iii) $l(p_2) \leq L$ and $l(p_1) \geq R \cdot L$ and $\beta(p_1, p_2) \leq \Theta_2$.

A problem may appear when the so-filtered boundary has two or several consecutive negative curvature points, also known as "U" concavities. Such points are counted as one concavity.

The parts number of the object is then formalized as:

$$N_p = N + 1 \tag{2}$$

where N is the number of concavities, we add one so as to count the body part of the object.

We decided to most favor objects with six parts (see Fig. 4). We observed that this number corresponds approximately to the number of significant parts of many common multi-part objects at a natural observation scale: hand, body, airplane, star, chair, fish, tools, etc. Moreover, Biederman [1] showed that human perceptual vision arrives to identify an object with two or three parts. If we want to take into account the possible partial self-occlusions of 3D multi-part objects, then six visible parts seems to be a good number.

The parts number criterion is computed as:

$$C_1 = f(N_p) \tag{3}$$

where $f(N_p)$ is a linear function (see Fig. 4).

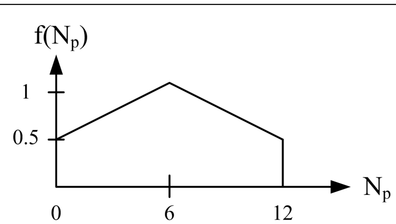


Figure 4. Part function.

In Fig. 5, once filtered, the hand boundary has several "U" concavities. Our approach permits to calculate six parts for this object, the same number as for the ideal case of the star.

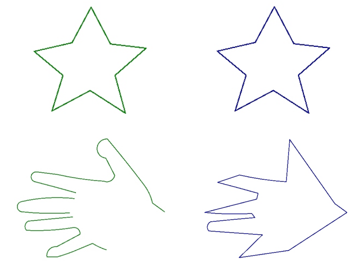


Figure 5. Parts number: (Left) Original images, (Right) Filtered images. The two objects have six parts.

3.1.2. Closure

$$C_2 = \frac{L_{1,n}}{L_{1,n} + G_{1,n}} \tag{4}$$

where: $L_{1,n} = \sum_{k=1}^{n} l(p_i)$ and $L_{1,n} = \sum_{k=1}^{n} l(h_i)$. $l(p_i)$ is the length of a primitive p_i and $l(h_i)$ the length of a gap h_i between two consecutive primitives.

3.1.3. Visual balance

There are different sorts of visual balance.

- Shape visual balance:

$$C_3 = 1 - \frac{dist(G, G_{bb})}{0.5D_1} \tag{5}$$

  $G = \frac{1}{n}\sum_{i=1}^{n} g_i$, and G_{bb} are the centroids of the subset and of the bounding box respectively, D_1 is the diagonal length of the bounding box. g_i is the mass center of a primitive; for a line segment, it is its middle point, but for an arc segment, it is computed as:

$$g_{i_x} = R\frac{sin\alpha cos\theta}{\alpha} + C_x$$
$$g_{i_y} = R\frac{sin\alpha sin\theta}{\alpha} + C_y$$

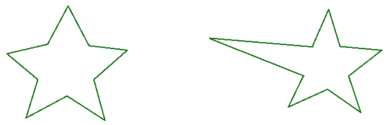
  where $\alpha = \frac{\gamma}{2}$, $\theta = \phi + \frac{\gamma}{2}$, ϕ and γ are the starting and the sweeping angle of the arc respectively, R is the arc radius, and (C_x, C_y) the arc center.

  An example is presented in Fig. 6 where the first case looks better balanced; this is attested by a higher value of C_3 than for the second case.

- Location visual balance: A subset is perceptually salient if it is located at the center of the scene.

$$C_4 = 1 - \frac{dist(G, G_I)}{0.5D} \tag{6}$$

  where G and G_I are the centroids of the subset and of the image, respectively, D is the diagonal length of the image, and $dist()$ is the distance between two points.



(a) (b)

Figure 6. Shape visual balance: (a) 0.92 , (b) 0.79.

- Size visual balance: A large object in the image is more salient than a small one.

$$C_5 = 1 - \frac{A - A_s}{A} \quad (7)$$

where A and A_s are the area of the image and of the subset respectively.

3.1.4. Smoothness

$$C_6 = \frac{1}{n} \sum_{i=1}^{n} \frac{1 + \cos \beta_i}{2} \quad (8)$$

where β_i is the turn angle between two consecutive oriented primitives, n is the number of primitives of the subset.

An example is shown in Fig. 7, where the second subset is a less regular shape than the first one. This is validated by a smaller value of C_6 for (b).

(a) (b)

Figure 7. Smoothness: (a) 0.80 , (b) 0.67.

3.1.5. Gap distribution

A subset seems to have better perceptual significance if the gaps are equally distributed on its boundary than if they are random. If there are no gaps, this criteria is equal to 1, otherwise, it is computed as:

$$C_7 = 1 - \frac{dist(G_r, G_v)}{0.5 D_1} \quad (9)$$

where G_r and G_v are the mass centers of the real and virtual primitives, as computed before.

In Fig. 8, by visual observations, the subset (b) obtained higher subjective confidence than the subset (d), which is validated by higher formal value of C_7.

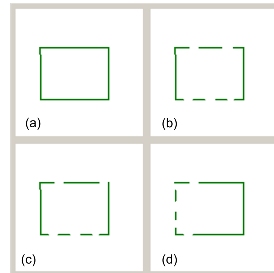


Figure 8. Gap distribution (a) 1, (b) 0.90, (c) 0.85, (d) 0.20.

3.1.6. Compactness

A good shape must be relatively compact. A circle is the most compact form, but as we are looking for multi-part objects, we propose:

$$C_8 = 1 - \frac{4\pi A}{P^2} \quad (10)$$

where P is the perimeter and A is the area of the subset.

3.1.7. Border Criterion

This criterion, noted C_9, has just a boolean value: if the subset bounding box is too close to one of the image borders, it is equal to 0, otherwise it is equal to 1.

3.1.8. Criteria combination

In order to compute the formal main score presented in Section 2, we have to integrate the global criteria C_k into the single measure $v(S_i)$.

One of the well-known multi-criteria methods is the additive combination of those criteria weighted by parameters to be adjusted:

$$v(S_i) = \sum_{k=1}^{K} w_k(S_i) C_k(S_i) \quad (11)$$

where K is the number of criteria.

We define that $v(S_i)$ must decrease as perceptual significance of a valid subset S_i decreases. A valid subset is one that respects the constraints of non-crossing and simple cycle of primitives.

3.2. Results for multi-part objects detection

In this section, we present the results of SAFE methodology on five images (see Fig. 9).

3.2.1. Validation measures

The subjective validation (SV) is evaluated by a recall estimate, noted R:

$$R = \frac{1}{n} \sum_{i=1}^{n} P_i \quad (12)$$

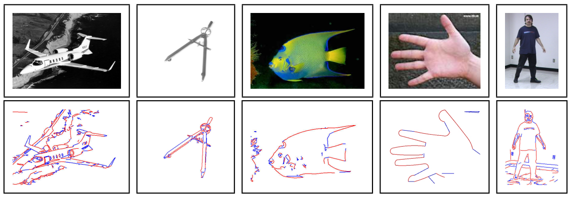


Figure 9. Original images and their corresponding primitive maps.

with

$$P_i = \begin{cases} 1 & \text{if } b(S_i) = a(S_i) \\ 0 & \text{else} \end{cases}$$

R defines the proportion of validated subsets, i.e. those such that $b(S_i) = a(S_i)$.

In order to evaluate the discrepancy of the ranking within the subsets list, we compute a precision estimate, noted P:

$$P = 1 - \frac{1}{n}\sum_{i=1}^{n} r_i \qquad (13)$$

with:

$$r_i = \frac{|b(S_i) - a(S_i)|}{M-1}$$

r_i is the normalized difference between the ranks, M is the total number of subsets, i.e. the worst rank difference between $b(S_i)$ and $a(S_i)$ is equal to $(M-1)$, and n is the number of subsets for which we want to compute R and P.

Note that these definitions are similar but not identical to the ones used in content-base image retrieval. Here, they estimate the rank similarity between the subjective and formal sets S and F.

3.2.2. Experiments

A subjective set S corresponding to possible solutions is built up (see Section 2), for each image in Fig. 9 (SGT).

When forming the list of solutions, it is preferable that it contains a large variety of situations, so that a great number of criteria can be investigated to model the human perception of a good shape, for a given category. However, generating such a large list of possible solutions is not an easy task since there could be a huge combination of primitives. For example, we have produced a list of 1000 random subsets, for a given image. The problem is that this so-built list does not illustrate the variety of situations, that we, as human being, can predict. Therefore, this experiment of a large random list was not conclusive. Thus, we opted for a manual selection so that we can choose the subsets in accordance to predictable situations like occlusions, background clutter, gaps and so on.

We have observed that the production of a large subjective set S is a challenging task. As pointed out in [1], accurate absolute judgment of human visual system rarely exceeds 7 ± 2 categories and comparing reliably a huge number of subsets with too small differences is not easy. Therefore, we decided to work with a reduced list of subsets, more exactly 30 subsets for each image, but that have large discrepancy between them.

The formal main score $v(S_i)$ is computed for each subset by incorporating, one by one, each formal global criterion (FGT). The ranking evolution between the formal set and the subjective one is observed, i.e. $b(S_i)$ is compared to $a(S_i)$ (SV). The recall values for each criterion except for C_1, C_4 and C_9, are presented on Table 1. The recall values for these latter criteria are not presented since they are similar for several subsets and do not allow the sorting. The obtained results are very poor, i.e. a unique Gestalt cue is not sufficient to reliably detect an object. Nevertheless, for the angel fish image, the result is not too bad since 50% of the top ten formal subsets obtained the same subjective judgment.

Images	C_2	C_3	C_5	C_6	C_7	C_8
Airplane	10%	10%	10%	0%	10%	0%
Compass	0%	0%	0%	0%	0%	0%
Angel fish	50%	0%	20%	0%	10%	0%
Hand	0%	0%	20%	0%	0%	0%
Man	0%	0 %	20%	10%	0%	0%

Table 1. Recall for each single criterion.

In Fig. 10(a) are presented the first nine subjective subsets for the airplane image, the best subjective solution is on the top left. The subjective rank is displayed on the top left of each subset. In Fig. 10(b) are shown the first nine subsets sorted in accordance to the closure criterion, which was agreed as powerful cue by several works [2], [10], [14]. We observe that bad solutions, in accordance to human judgment, are ranked among the best ten solutions of the formal set F, e.g. the subset $\#19$ or $\#18$ have low subjective scores.

In this paper, we determined the weights of the linear multi-criteria combination empirically. For the set of images presented in Fig. 9, we have used the following parameter values:

- combination weights: $w_k = 1$, except for C_2, C_5 and C_6 for which $w_k = 5$.
- filtering: $\Theta_1 = 15$, $\Theta_2 = 40$, $L = 40$, $R = 10$.

The result obtained with the multi-criteria formal main score $v(S_i)$, for the airplane image is shown in Fig. 10 (c) and (d). The top nine best (c) and nine worst subsets (d) of F are presented. The best formal solution is on the top left of Fig. 10(c).

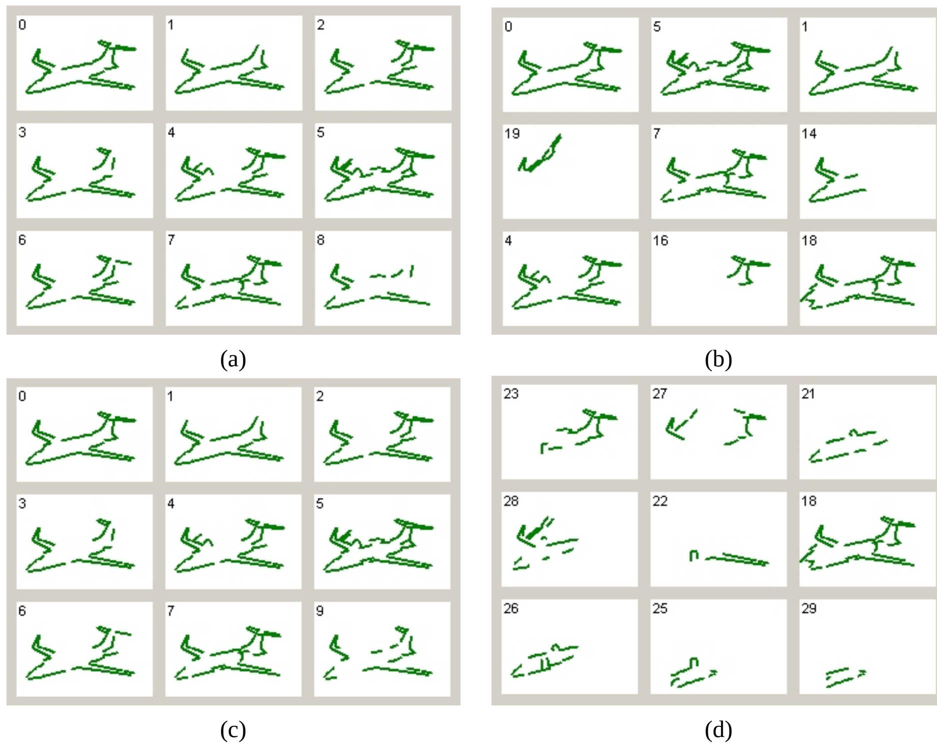


Figure 10. (a) Top nine subjective subsets, (b) Top nine subsets sorted by closure criterion, (c) Nine best formal subsets, (d) Nine worst formal subsets. Recall = 80%, Precision = 93%.

We can note on this figure that the best perceptually significant subsets are among the best formal ones, and the worst ones are really among the worst formal ones. Note also the subset $\#18$, which was ranked among the top ten best subsets for the single closure criterion. Thanks to the multi-criteria combination, this subset is ranked among the ten worst, which is coherent with its subjective ranking.

Similar improvements are obtained with the other multi-part objects. For example, the formal result obtained with the multi-criteria combination for the man image is shown in Fig. 11. The top nine best (a) and nine worst subsets (b) of F are presented. The best formal solution is on the top left of Fig. 11(a).

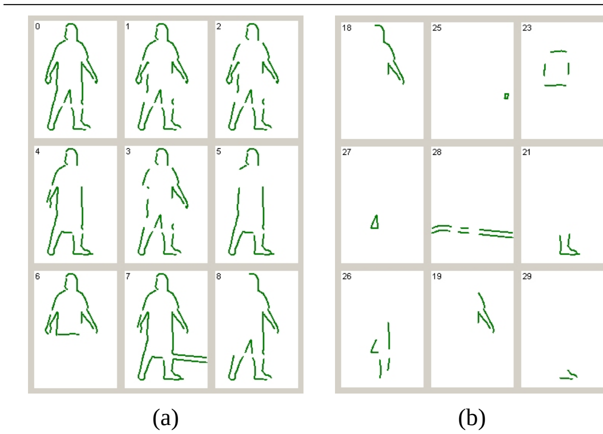


Figure 11. (a) Nine best formal subsets, (b) Nine worst formal subsets. Recall = 80%, Precision = 93%.

We have observed that it is more interesting to compute the recall value for the top ten best and the ten worst subsets because these two sets reflect more the reliability of the main global score $v(S_i)$. For the precision, we take the whole list of subsets because this value is normalized to the formal and subjective sets cardinality. The recall and precision values for the images in Fig. 9 are presented on Table 2.

The multi-criteria combination improves considerably the recall values, and the precision values are quite high. This indicates that the formal evaluation of these solutions tends to be similar to the human judgment. This important result is needed for computer vision systems as Jain *et al.* mentioned in their paper [6]. Indeed, they observed, for their shape-based trademarks retrieval system, that the ranking of the top ten images retrieved by their artificial system is significantly different from the ranking of those selected by humans. To their point of view, this clearly shows that com-

Images	*Recall*	*Precision*
Airplane	80%	93%
Compass	70%	88%
Angel fish	70%	93%
Hand	60%	93%
Man	80%	93%

Table 2. Recall and precision for the first top ten subsets of the images in Fig. 9.

puter vision systems should take into account human visual perception to design segmentation algorithms. The SAFE methodology tends toward achieving this goal.

4. Conclusion and Future Work

We have proposed an approach that combines top-down processing and perceptual grouping in order to detect a particular category of objects, in this case multi-part objects. Instead of being model-based as existing approaches in the literature, it is based on the identification of a set of global criteria that formalize the abstract conceptual category of multi-part objects. An original methodology, called SAFE, is presented that permits to identify and validate the formal model of these global criteria by human psycho-evaluation. These criteria are in accordance with the law of Prägnanz, which represents the global percept of a grouping.

The detection of these so-categorized objects is not easy since large and complex groupings must be generated. The extraction of a perceptually significant grouping is a difficult problem but the perceptual evaluation of this grouping is even more challenging as it entails the visual assessment of its proper shape. We address this latter problem and evaluate the quality of an hypothesized multi-part object boundary.

The proposed approach is designed to detect a generic category of objects, but it could be further refined to deal with the categorization of specific objects. The conceptualization of an object category is a difficult task, therefore great efforts must be done in order to determine additional criteria. Linear combination of these criteria is the most straightforward method, but since reliable weights are needed in order to control the balance between each criterion, other methods must be investigated. The next step of our work will be to automatically segment an image so that primitives that optimize a multi-criteria cost function are extracted. We will investigate more specifically graph-theoretic approaches.

The preliminary results presented in this paper are encouraging since they demonstrate that it is possible to design artificial visual systems which tend to perform as well as humans. Though important works have been conducted, existing computer vision systems are still far from attaining the human visual performance. The original methodology presented in this paper is a step toward this.

Acknowledgments This work is supported by a discovery grant from Science and Engineering Research Canada (NSERC).

References

[1] I. Biederman. Recognition by components: A theory of human image understanding. *Psychological Review*, 94(2):115–147, 1987.

[2] J. Elder and S. Zucker. Computing contour closure. In *Proc. 4th European Conference on Computer Vision*, pages I:399–412, 1996.

[3] F. Estrada and A. Jepson. Perceptual grouping for contour extraction. In *17th International Conference on (ICPR'04)*, pages II: 32–35, Cambridge UK, August 23 - 26, 2004.

[4] Q. Iqbal and J. Aggarwal. Retrieval by classification of images containing large man-made objects using perceptual grouping. *Pattern Recognition*, 35(7):1463–1479, July 2002.

[5] D. Jacobs. Robust and efficient detection of salient convex groups. *IEEE Trans. Pattern Anal. Machine Intell.*, 18:23–37, 1996.

[6] A. Jain and A. Vailaya. Shape-based retrieval: A case-study with trademark image databases. *Pattern Recognition*, 31(9):1369–1390, September 1998.

[7] K. Kohler. *Gestalt Psychology*. New York:Liveright, 1929.

[8] L. Latecki and R. Lakamper. Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding*, 73(3):441–454, March 1999.

[9] D. Lowe. *Perceptual Organization And Visual Recognition*. Kluwer Academic Publishers,Boston, Mass., 1985.

[10] S. Mahamud, L. Williams, K. Thornber, and K. Xu. Segmentation of multiple salient closed contours from real images. *IEEE Trans. Pattern Anal. Machine Intell.*, 25(4):433–444, April 2003.

[11] R. Mohan and R. Nevatia. Perceptual organization for scene segmentation and description. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(6):616–635, June 1992.

[12] S. Sarkar and K. Boyer. Perceptual organization in computer vision: A review and a proposal for a classificatory structure. *IEEE Trans. Systems. Man, Cybernet.*, 23(2):382–399, 1993.

[13] S. Sarkar and P. Soundararajan. Supervised learning of large perceptual organization: Graph spectral partitioning and learning automata. *IEEE Trans. Pattern Anal. Machine Intell.*, 22(5):504–525, May 2000.

[14] S. Wang, K. T., and S. J. M. Salient boundary detection using ratio contour. In *Neural Information Processing Systems Conference (NIPS)*, pages 1571–1578, Vancouver, Canada, 2003.