PERCEPTUAL GROUPING VIA UNTANGLING GESTALT PRINCIPLES

Yonggang Qi[†] *Jun Guo*[†] *Yi Li*[⋆] *Honggang Zhang*[†] *Tao Xiang*[⋆] *Yi-Zhe Song*[⋆] *Zheng-Hua Tan*[‡]

[†] School of Information and Communication Engineering, BUPT, Beijing, China
[⋆] School of EECS, Queen Mary, University of London, London E1 4NS, U.K.
[‡] Department of Electronic Systems, Aalborg University, Aalborg, Denmark

ABSTRACT

Gestalt principles, a set of conjoining rules derived from human visual studies, have been known to play an important role in computer vision. Many applications such as image segmentation, contour grouping and scene understanding often rely on such rules to work. However, the problem of Gestalt confliction, i.e., the relative importance of each rule compared with another, remains unsolved. In this paper, we investigate the problem of perceptual grouping by quantifying the confliction among three commonly used rules: similarity, continuity and proximity. More specifically, we propose to quantify the importance of Gestalt rules by solving a learning to rank problem, and formulate a multi-label graph-cuts algorithm to group image primitives while taking into account the learned Gestalt confliction. Our experiment results confirm the existence of Gestalt confliction in perceptual grouping and demonstrate an improved performance when such a confliction is accounted for via the proposed grouping algorithm. Finally, a novel cross domain image classification method is proposed by exploiting perceptual grouping as representation.

***Index Terms*—** Gestalt confliction, RankSVM

1. INTRODUCTION

Human visual system is very powerful so that we can easily find sense from chaos. In neuroscience, how the brain generates or perceives visual objects is a critical problem. Perceptual grouping holds the concept that humans perceive some elements of the visual field as going together more strongly than the others. Thus individual objects are formed.

The Gestalt psychologists define perceptual grouping as grouping the components of an visual scene in a proper way so as to form an individual visual object. Wertheimer [1], a pioneer in the Gestalt school, pointed out the significance of perceptual grouping and further listed several key factors, such as proximity, similarity and continuation. His work has triggered a large amount of research aimed at understanding human visual systems [2, 3].

However, the problem of Gestalt confliction, i.e., how Gestalt principles interact with each other and which plays a more important role than the others, remains open [4]: Although every proposed single Gestalt principle has been proven to be useful for grouping in computer vision when used alone [5, 6, 7], very few work attempts to investigate how they can be exploited jointly in a single framework. This problem has been the subject of investigation in the fields of psychology and psychophysics. However, even psychologists know very little about how exactly they interact in human vision system [8], thus shedding little light on how to design a computer vision system. Recently, the problem of Gestalt confliction is tackled in computer graphics. Nan et al. [9] simplifies architectural drawings with conjoining gestalt rules. Although achieving good performance on drawings, the method cannot be applied to natural images, which is a much more challenging task.

Our goal in this paper is to formulate a general framework to learn the relative importance among different Gestalt principles explicitly for a given grouping task.We particularly focus on the line segments grouping problem although the framework can be applied to any other grouping task. In line segment grouping, only the geometric information is used. Solving this problem thus can provide a bridge to relate different types of images, e.g. natural images vs human sketches. An example is shown in Fig.1. It can be seen that while it is difficult to match a natural image of a beer-mug with a sketch image of the same object directly, matching them becomes much easier if we could convert the natural image into a sketch. To this end, one can assume that an object sketch equals to a collection of the most salient edges of the object in a natural image. One thus can automatically generate a sketch out of a natural image by grouping edges in a natural image, followed by filtering out noise edges. In this paper, after solving the grouping problem by exploiting Gestalt confiliction, we further propose to utilize the automatically generated sketches for cross domain object classification.

To this end, a new data set is created by collecting a set of training images from the Berkeley Segmentation Data Set 500 (BSDS500) [10], in which the boundaries of each image are drawn manually. Line segments are then extracted on the basis of the drawn boundaries. Therefore, line segments can replace the boundaries of each image to represent the object contours. Then we assign each line segment with a label manually, and such labeled line segments will serve as ground

Fig. 1. Beer-mugs: **Left**.natural image. **Middle**.human drawn sketch. **Right**. automatically generated sketch.

truth. For solving the Gestalt confiliction problem, a learning to rank strategy based on RankSVM [11] is proposed to learn the relative importance among three Gestalt principles, namely proximity, continuity and similarity. A multi-label graph-cuts [12, 13, 14] perceptual grouping framework is further developed to group contour line segments by using the learned importance of different Gestalt principles. Finally, a cross domain image classification task is tackled based on this perceptual grouping framework.

This paper has the following contributions: (1) A learning strategy based on RankSVM is proposed to learn the relative importance among three Gestalt principles, namely proximity, continuity and similarity. Note that the proposed framework is very general that it can be used to study relative importance of other Gestalt principles. (2) A new data set is presented which contains 100 images selected from the BSDS500 perceptual grouping benchmarking dataset. In each image the extracted line segments are manually labeled by human for both model training and evaluation. (3) A novel application based on graph-cuts is proposed for perceptual grouping by conjoining Gestalt principles, while addressing the problem of Gestalt confliction. More specifically, we cast the perceptual grouping in natural images as a line grouping problem and explicitly define three Gestalt principles involving geometry information among lines only, motivated by Song et al. [15]. (4) For the first time, the perceptual grouping is used for cross domain image classification.

2. LEARNING GESTALT CONFLICTION

2.1. Data Set

In order to learn the relative importance of three Gestalt principles, namely similarity, continuity and proximity, we propose a new data set which contains 100 images from BSDS500 [10], where line segments in each image are labeled manually into groups of semantic objects by using the three principles.

More specifically, we first perform contour extraction by a state-of-the-art algorithm proposed in [10] to get the Ultrametric Contour Map (UCM) of an image. We then further simplify the UCM by using the Line Segment Detector (LSD) [16], which converts continuous boundary lines to disjoint line segments. Such line segments are derived from psychological concept about how humans perform the same task.

Fig. 2. Examples of Data Set

Then we assign labels to these line segments Q manually. A number of examples are shown in Fig.2, where line groups are color coded.

In this process, we essentially throw away the photometric properties. This choice of primitives is motivated by the work in [17], which shows that sketching, despite throwing away photometric properties such as color, is a universal form of communication. People describe the visual world with sketch-like petroglyphs or cave paintings since prehistoric times. And the ability to draw and recognize sketched objects is ubiquitous. Recent neuroscience work [18] also indicates that simple, abstracted sketches activate our brain in similar ways to real stimuli.

2.2. Learning Gestalt Confliction

Given the training images, the problem of learning the importance of different gestalt principles is cast into a learning to rank problem. Learning to rank has been studied extensively in the field of document retrieval. In our problem, for any line segment we treat it as a query and go through all the other line segments in the image and rank them (retrieving) according to whether they belong to the same group, that is, rank 1 is assigned if they are, and rank 2 otherwise. In essence, the ranking model is to learn a weighted distance using the similarity measured by the three principles, so that this ranking order is maintained as much as possible across all training images.

The RankSVM model is adopted which has been used in computer vision problems such as person re-identification and gait recognition [19]. Formally, in our case, the training set is comprised of:

- **Set of line segments**. $Q = \{q_1, q_2, \ldots, q_{|Q|}\}$, where $|Q|$ is the number of line segments in Q.
- **Pair of line segments**. Each pair (q_i, q_j) is described by a 3D feature vector $\mathbf{x}(q_i, q_j)$, which shows the distance or difference of the pair of line segments. Each

dimension in $\mathbf{x}(q_i, q_j)$ corresponds to a Gestalt principle. More specifically, $\mathbf{x}(q_i, q_j)$ is defined as:

$$\mathbf{x}(q_i, q_j) = \begin{bmatrix} x(q_i, q_j)_{sim} \\ x(q_i, q_j)_{con} \\ x(q_i, q_j)_{pro} \end{bmatrix}$$

where

$$x(q_i, q_j)_{sim} = \frac{R_l\{q_i, q_j\} + R_s\{q_i, q_j\}}{2}$$

$$x(q_i, q_j)_{con} = R_s\{q_i, q_j\}$$

$$x(q_i, q_j)_{pro} = R_p\{q_i, q_j\}$$

Where R_l, R_s and R_p are the relationships of length, slope and distance between pair of line segments respectively. Essentially, $x(q_i, q_j)_{sim}$ indicates how similar the two line segments are by measuring the relationship of both length and slope between q_i and q_j. Similarly, $x(q_i, q_j)_{con}$ measures the continuity between q_i and q_j, and $x(q_i, q_j)_{pro}$ measures the distance between q_i and q_j.

- **Relationship between every pair**. Each line segment q_i is labeled by a relevance indicator $y(q_i, q_j)$ which represents its relationship to another line segment q_j. In our case, we define y with a value 1 when a line segment q_i is grouped together with another line segment q_j, and -1 otherwise. Thus, for each line segment, we divide all the other line segments into two sets depending on its relevance indicator with q_i:

$$Q(q_i)^+ = \{{q_1}^+, {q_2}^+, \ldots, q_{|Q(q_i)^+|}\}$$

where $y(q_i, {q_j}^+) = 1$ for all ${q_j}^+ \in Q(q_i)^+$, similarly,

$$Q(q_i)^- = \{{q_1}^-, {q_2}^-, \ldots, q_{|Q(q_i)^-|}\}$$

where $y(q_i, {q_j}^-) = -1$ for all ${q_j}^- \in Q(q_i)^-$.

It will form the positive pairs $\hat{Q}^+ = (q_i, {q_j}^+)$, and the relative negative pairs $\hat{Q}^- = (q_i, {q_j}^-)$. Thus, we obtain preference pairs $P = (\hat{Q}^+, \hat{Q}^-)$.

With the constraints P, we can learn the ranking function, $f(q_i, q_j) = \boldsymbol{\omega}^T \mathbf{x}(q_i, q_j)$, where $\boldsymbol{\omega}$ refers to a 3D weight vector which indicates the significance of each Gestalt principle in grouping. Specifically, we obtain $\boldsymbol{\omega}$ in the learning function by solving the following optimization problem:

$$\omega = argmin_{\omega} \frac{1}{2} ||\omega||^2 + C \sum_{k=1}^{|P|} l(\omega^T (\hat{Q}^+ - \hat{Q}^-)) \qquad (1)$$

where k is the index of the preference pairs, $|P|$ is the total number of preference pairs used for training, C is a positive importance weight on the ranking performance and is automatically selected by cross validation on the training set. l is the hinge loss function.

	similarity	continuity	proximity
α	0.425	0.436	0.139

Table 1. Inverse importance of gestalt rules learned by RankSVM. Proximity with the smallest weight, hence it is the most important principle.

2.3. The Learned Importance

Although the efficient primal RankSVM algorithm is adopted, the amount of training data is still too big for the model to be tractable. To solve this problem, we sub-sample the data from each training data. Specifically, in each image, we randomly choose two line segments in every group, and then all the selected line segments are used to form the positive pairs and relative negative pairs. Finally, approximate five million preference pairs are formed for learning the RankSVM model.

We use α, which derives from $\boldsymbol{\omega}$, to indicate the learned importance. More specifically, α is given by normalizing each dimension of $\boldsymbol{\omega}$ to 0-1. Therefore, α has three values corresponds to each of the three Gestalt principles. Note that the value of α indicates the inverse importance of each Gestalt principle.

Table 1 shows the learned inverse importance value for the three Gestalt principles using our RankSVM model (More details can be found in settings in section 4). Note that smaller value corresponds to higher importance. This table shows clearly that the proximity principle is the most important one compared to the other two. This finding is in tune with the results of psychology study [20] which suggests that humans also rely more on the proximity principle than the others. Then the learned relative importance is used to fuse the three Gestalt principles in a graph-cuts based model as proposed in the next section.

3. A MULTI-LABEL GRAPH-CUTS MODEL FOR GROUPING LINE SEGMENTS

In this section, a multi-label graph-cuts [13] based model for grouping line segments is proposed to validate the effectiveness of the learned relative importance of the three Gestalt principles. The problem of grouping line segments is treated as a min-cut/max-flow optimization problem. Fig.3 gives an overview of our proposed perceptual grouping framework.

3.1. Potential Groupings

The grouping primitives are line segments, which are the same as the ones used in training data, which is introduced in section 2. To solve this line segments grouping problem by the multi-label graph-cuts model, we need to specify the relationship between primitives and the possible labels, thus to compute the data cost item required in the subsequent

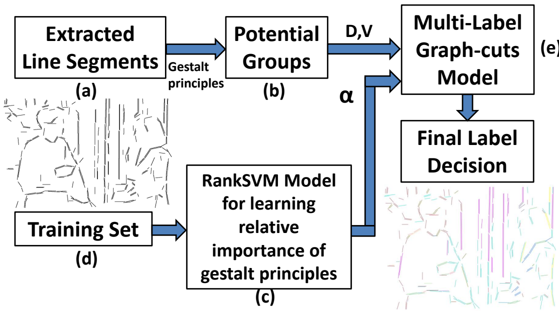

Fig. 3. **An overview of perceptual grouping framework**. First, line segments (a) are extracted as the primitives. Potential groups (b) are then defined according to three popular gestalt principles. Three principles are combined in a Multi-label Graph-cuts model (e). In this model, the gestalt principles are given different importance value α, learned using a RankSVM model (c) on a training set (d). Finally, the optimal grouping is obtained by the multi-label graph-cuts model (e).

multi-label graph-cuts framework. The possible labels are obtained by discovering potential groupings, which are sets of primitives, indicating the possibility to assign a label to one primitive.

In our case, potential groupings are defined by three Gestalt rules: similarity, continuity and proximity. Given a set of line segments , Q , which consists of n line segments in an image, each line segment $q_i \in Q$ will in turn serve as one where all other line segments $q_j \in Q$ are compared against with. More specifically, each set of potential groupings is defined as follows.

Similarity Gestalt is defined by detecting line groups which share high similarity of both length and slope. A similarity group is defined as:

$$L_i^{sim} = \bigcup \{q_i, q_j\} | \{R_l\{q_i, q_j\} \ > \ t_l \ \& \ R_s\{q_i, q_j\} > t_s\}$$

Thus the set of similarity groups is $L^{sim} = \bigcup\limits_{i=1,2,\ldots,n} L_i^{sim}$.

Continuity Gestalt is defined by detecting line groups which would form a continuous curve.

$$L_i^{con} = \bigcup \{q_i, q_j\} | \{R_s\{q_i, q_j\} \ > \ t_s\}$$

Hence the set of continuity groups is $L^{con} = \bigcup\limits_{i=1,2,\ldots,n} L_i^{con}$.

Proximity Gestalt is defined by detecting line groups where line segments are close enough to each other.

$$L_i^{pro} = \bigcup \{q_i, q_j\} | \{R_p\{q_i, q_j\} \ > \ t_p\}$$

Similarly, the set of proximity groups is $L^{pro} = \bigcup\limits_{i=1,2,\ldots,n} L_i^{pro}$.

t_s, t_l and t_p in the above equations are fixed thresholds for determining whether a pair of line segments should be grouped into a potential group or not when applying one of three Gestalt principles as grouping criterion in turn.

Given the potential groupings, data cost item, which will be detailed in the next section, is naturally obtained.

3.2. Multi-Label Graph-Cuts Model

The problem of grouping line segments is formulated as a min-cut/max-flow optimization problem. And the overall energy function is defined as:

$$E(L) = \sum_{q_i \in Q} \alpha D(q_i, L) + \sum_{\{q_i, q_j\} \in N} V_{\{q_i, q_j\}} \tag{2}$$

where N is the set of pairs of neighboring elements in Q, $L = \{L^{sim}, L^{con}, L^{pro}\}$, and α is the learned relative importance given by RankSVM. D is data cost energy and V is smoothness cost energy, these two items are detailed in the following.

Data cost item measures the fitness between the line segment q_i and the possible assigned label L_i. The higher the fitness, the lower the cost or penalty. More specifically, in our case:

A *similarity data cost* between q_i and a possible label L_i^{sim} is defined as:

$$D(q_i, L_i^{sim}) = 1 - \frac{1}{|L_i^{sim}|} \sum_{q_j \in L_i^{sim}} \frac{\{R_l\{q_i, q_j\} + R_s\{q_i, q_j\}\}}{2}$$

where L_i^{sim} (i=1,2,...,n) is a label, which is used to represent a potential grouping found by similarity Gestalt principle. $|L_i^{sim}|$ is the number of pairs of line segments in this potential grouping. Essentially, we can see from this equation that the better the line segments q_i and q_j obeys similarity principle, the more possible q_i is assigned with label L_i^{sim}. Similarly, continuity and proximity data cost will be defined as follows.

A *continuity data cost* between q_i and L_i^{con}:

$$D(q_i, L_i^{con}) = 1 - \frac{1}{|L_i^{con}|} \sum_{q_j \in L_i^{con}} R_s\{q_i, q_j\}$$

A *proximity data cost* between q_i and L_i^{pro}:

$$D(q_i, L_i^{pro}) = 1 - \frac{1}{|L_i^{pro}|} \sum_{q_j \in L_i^{pro}} R_p\{q_i, q_j\}$$

Smoothness cost item measures the spatial correlation between neighboring elements. Elements with a smaller distance have higher probability of belonging to the same gestalt group. Between two neighboring elements q_i and q_j, the smoothness energy is defined by the inverse Euclidean Hausdorff-distance between them, which is the same to the one used in [9].

$$V_{\{q_i, q_j\}} = d(q_i, q_j)^{-1}$$

To this end, upon solving above optimization problem in eq (2), each line segment will be assigned with a optimal label.

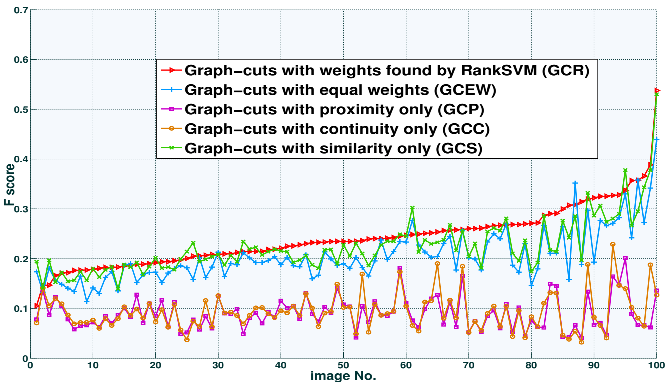

Fig. 4. Grouping performance comparison. Note that F scores are sorted according to GCR to show in figure.

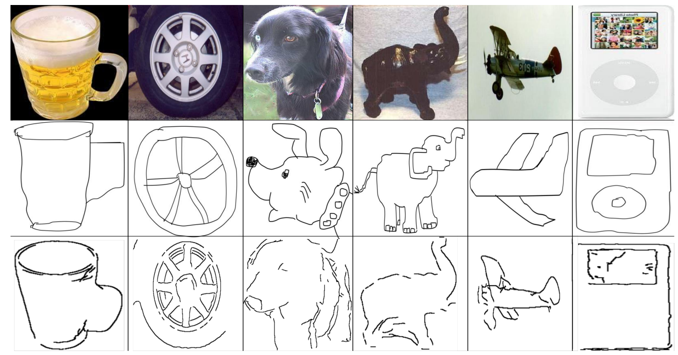

Fig. 5. From Top to Bottom: Example natural images, human drawn sketches and our generated sketches. We can see that our generated sketches are simple, basic and similar to human drawn sketch.

4. RESULTS AND ANALYSIS

Settings – Leave-one-out cross validation is adopted in our experiments: 100 images are divided into two parts with 99 images as the training set for learning the importance of different principles, and the left one for testing for the grouping evaluation. This is repeated 100 times. And α shown in Table 1 is averaged over 100 trials.

Evaluation metrics – F-measure is used to evaluate our line segments grouping performances. For each test image with n line segments, the grouping obtained using our graph-cuts algorithm is represented as an affinity matrix. Given the ground truth (human grouping result), another affinity matrix is constructed. Using these two matrices we compute the F-measure to evaluate how well the estimated grouping matches against the ground truth grouping.

Grouping results and discussions – We compare the grouping result obtained using our algorithm with the learned importance weight against those by using the same graph-cuts algorithm but either (1) uses an equal weight to the three principles (GCEW), or (2) uses one of the three principles alone. The result on each of the 100 images is shown in Fig. 4. It shows clearly that for all 100 images, our algorithm with the learned weighting consistently outperforms the algorithm with equal weight assigned to the three principles. On average, an increase of 3.98% in the F-measure score is obtained. This result demonstrates that the learned weighting not only supports the psychology study findings, but also has practical use in solving computer vision problems. Also, we can see that the performance of our algorithm is better than any gestalt principle is used alone. In particular, it is interesting to note that when a single gestalt principle is used, similarity is the best option; however, when naively combined (i.e. giving equal weight), the result is even worse than using similarity alone. Our learned weighting suggests that a larger weight should be given to proximity if the combination is to yield any improvement on the grouping performance.

5. CROSS DOMAIN IMAGE CLASSIFICATION

In this section, assuming the learned relative importance is general, we apply it to a simple cross domain image classification application on Caltech256 [21] where 480 images are collected from 6 categories with 80 images in each category. To obtain sketches from natural images, continuous curve segments instead of line segments are extracted from natural images and serve as grouping primitives for perceptual grouping as proposed in section 3. Given the grouping results of continuous curve segments, sketches are obtained by simply filtering away the curves of groups whose total length is less than a fixed value[1]. A number of sketch examples are shown in Fig.5.

Furthermore, to evaluate how well these automatically generated sketches can be recognized by computer, SVM classifiers are trained using human drawn sketches from large scale human drawn sketch dataset [17], which consists of 250 categories and 80 sketches in each category. Therefore, 250 classifiers are trained, and then used to classify these automatically generated sketches. **Canny**, **Pb** [10] edge detector produced sketches and Primal sketch (**PS**) [22] serve as the competitors. Table 2 shows the classification results in which our sketches outperform the state-of-the-art alternatives.

	Canny	**Pb**	**PS**	**Ours**
Airplane	10.00	26.25	15.00	**28.75**
Car-tire	15.00	**23.75**	8.75	22.50
Elephant	0	2.50	1.25	**10.00**
Ipod	13.75	15.00	15.00	**15.00**
Beer-mug	15.00	**35.00**	15.00	31.25
Dog	0	3.75	3.75	**17.50**
Average	8.96	17.71	9.79	**20.83**

Table 2. Rank 10 classification rate (%)

[1]More complicated filtering strategy can be adopted while it is not the focus of this paper.

6. CONCLUSION

To our best knowledge, it was the first time that a general learning framework was proposed to investigate the relative importance of diverse Gestalt principles when they were used together. We focused on a complex perceptual grouping task on natural images, and we developed an algorithm to congregate them according to this learned importance. In addition, a new dataset was proposed which makes possible to learn it. The results suggested that proximity plays a more dominant role than similarity and continuity which coincided with the findings of psychology studies. Moreover, we demonstrated that with the learned importance, different principles could be exploited jointly to improve the performance than using them alone or combining them naively. In the end, a simple and novel cross domain classification application illustrates the effectiveness of our proposal.

7. REFERENCES

[1] M. Wertheimer, *Laws of organization in perceptual forms*, Harcourt, Brace & Jovanovitch, London, 1938.

[2] Arnon Amir and Michael Lindenbaum, "A generic grouping algorithm and its quantitative analysis.," *TPAMI*, vol. 20, no. 2, pp. 168–185, 1998.

[3] Xiaofeng Ren and Jitendra Malik, "Learning a classification model for segmentation.," in *ICCV*. 2003, pp. 10–17, IEEE Computer Society.

[4] Johan Wagemans, James H. Elder, Michael Kubovy, Stephen E. Palmer, Mary A. Peterson, Manish Singh, and Rüdiger von der Heydt, "A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization.," *Psychological bulletin*, 2012.

[5] Joachim S. Stahl and Song Wang, "Edge grouping combining boundary and region information.," *TIP*, vol. 16, no. 10, pp. 2590–2606, 2007.

[6] G. Papari and N. Petkov, "Adaptive pseudo-dilation for gestalt edge grouping and contour detection," *TIP*, vol. 17, no. 10, pp. 1950–1962, 2008.

[7] Nagesh Adluru, Longin Jan Latecki, Rolf Lakmper, Thomas Young, Xiang Bai, and Ari D. Gross, "Contour grouping based on local symmetry.," in *ICCV*. 2007, pp. 1–8, IEEE.

[8] Michael Kubovy and Martin van den Berg, "The whole is equal to the sum of its parts: A probabilistic model of grouping by proximity and similarity in regular patterns," *Psychological Review*, 2008.

[9] Liangliang Nan, Andrei Sharf, Ke Xie, Tien-Tsin Wong, Oliver Deussen, Daniel Cohen-Or, and Baoquan Chen, "Conjoining gestalt rules for abstraction of architectural drawings.," *ACM Trans. Graph.*, vol. 30, no. 6, pp. 185, 2011.

[10] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik, "Contour detection and hierarchical image segmentation.," *TPAMI*, vol. 33, no. 5, pp. 898–916, 2011.

[11] Olivier Chapelle and S. Sathiya Keerthi, "Efficient algorithms for ranking with svms.," *Inf. Retr.*, vol. 13, no. 3, pp. 201–215, 2010.

[12] Yuri Boykov and Vladimir Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *TPAMI*, vol. 26, no. 9, pp. 1124–1137, 2004.

[13] Yuri Boykov, Olga Veksler, and Ramin Zabih, "Fast approximate energy minimization via graph cuts.," *TPAMI*, vol. 23, no. 11, pp. 1222–1239, 2001.

[14] Andrew Delong, Anton Osokin, Hossam N. Isack, and Yuri Boykov, "Fast approximate energy minimization with label costs.," *IJCV*, vol. 96, no. 1, pp. 1–27, 2012.

[15] Yi-Zhe Song, Bai Xiao, Peter M. Hall, and Liang Wang, "In search of perceptually salient groupings.," *TIP*, vol. 20, no. 4, pp. 935–947, 2011.

[16] Rafael Grompone von Gioi, Jrmie Jakubowicz, Jean-Michel Morel, and Gregory Randall, "Lsd: A fast line segment detector with a false detection control.," *TPAMI*, vol. 32, no. 4, pp. 722–732, 2010.

[17] Mathias Eitz, James Hays, and Marc Alexa, "How do humans sketch objects?," *ACM Trans. Graph. (Proc. SIGGRAPH)*, vol. 31, no. 4, pp. 44:1–44:10, 2012.

[18] Mathias Eitz, Kristian Hildebrand, Tamy Boubekeur, and Marc Alexa, "Sketch-based image retrieval: Benchmark and bag-of-features descriptors.," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 11, pp. 1624–1636, 2011.

[19] Ral Martn-Flez and Tao Xiang, "Gait recognition by ranking.," in *ECCV (1)*. 2012, vol. 7572 of *Lecture Notes in Computer Science*, pp. 328–341, Springer.

[20] J. H. Elder and R. M. Goldberg, "Ecological statistics of Gestalt laws for the perceptual organization of contours.," *J Vis*, vol. 2, no. 4, pp. 324–353, 2002.

[21] G. Griffin, A. Holub, and P. Perona, "The Caltech-256," Tech. Rep., California Institute of Technology, 2007.

[22] Cheng en Guo, Song Chun Zhu, and Ying Nian Wu, "Primal sketch: Integrating structure and texture," *CVIU*, vol. 106, no. 1, pp. 5–19, 2007.