Conjoining Gestalt Rules for Abstraction of Architectural Drawings

Liangliang Nan[1]   Andrei Sharf[2]   Ke Xie[1]   Tien-Tsin Wong[3]   Oliver Deussen[4]   Daniel Cohen-Or[5]   Baoquan Chen[1]

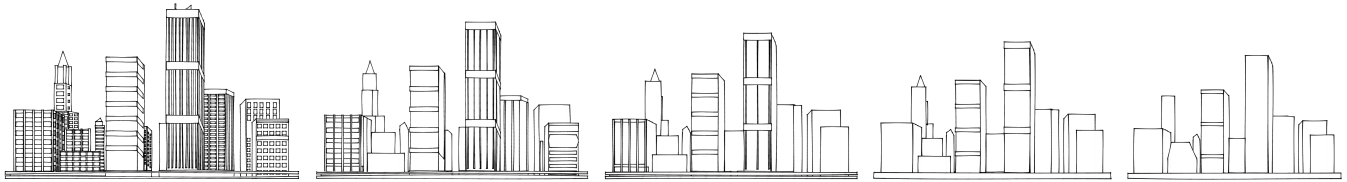[1] SIAT   [2] Ben Gurion Univ.   [3] CUHK   [4] Konstanz Univ.   [5] Tel Aviv Univ.

Figure 1: *Simplification of a complex cityscape line-drawing obtained using our Gestalt-based abstraction.*

Abstract

We present a method for structural summarization and abstraction of complex spatial arrangements found in architectural drawings. The method is based on the well-known Gestalt rules, which summarize how forms, patterns, and semantics are perceived by humans from bits and pieces of geometric information. Although defining a computational model for each rule alone has been extensively studied, modeling a conjoint of Gestalt rules remains a challenge. In this work, we develop a computational framework which models Gestalt rules and more importantly, their complex interactions. We apply *conjoining rules* to line drawings, to detect groups of objects and repetitions that conform to Gestalt principles. We summarize and abstract such groups in ways that maintain structural semantics by displaying only a reduced number of repeated elements, or by replacing them with simpler shapes. We show an application of our method to line drawings of architectural models of various styles, and the potential of extending the technique to other computer-generated illustrations, and three-dimensional models.

CR Categories: I.3.5 [Computer Graphics]: Geometric Modeling and Level of Detail—Geometric algorithms;

Keywords: simplification, abstraction, Gestalt

Links: DL PDF

1 Introduction

Artistic imagery, architectural renderings, cartography and games often exploit abstraction to clarify, exaggerate, simplify or emphasize the visual content. Abstraction is a strategy for communicating information effectively. It allows artists to highlight specific visual information and thereby direct the viewer to important aspects of the structure and organization of the scene. In this paper, we present an approach to the abstraction of 2D shapes, in particular

those of architectural models. Our approach to abstracting shape directly aims to clarify shape and preserve meaningful structures using Gestalt principles.

The well-known Gestalt principles by Wertheimer [1923], reflect strategies of the human visual system to group objects into forms and create internal representations for them. Whenever groups of visual element have one or several characteristics in common, they get grouped and form a new larger visual object - a *gestalt*. Psychologists have tried to simulate and model these principles, by finding computational means to predict what human perceive as gestalts in images.

The notion of Gestalt is very well-known and widely used in various fields. In particular, it explains the tendency of the human visual recognition to form whole shapes and forms just from bits and pieces of geometric information. Naturally, Gestalt principles have been used in computer vision, primarily in context with object recognition and scene understanding. In computer graphics, Gestalt principles have been applied to a variety of applications, like scene completion [Drori et al. 2003], image and scene abstraction [Wang et al. 2004; Mehra et al. 2009], stroke synthesis [Barla et al. 2006; Ijiri et al. 2008] and emerging images generation [Mitra et al. 2009]. In general, these works rely on discrete Gestalt principles, but none addresses the complex interactions emerging from the multitude of Gestalt principles operating simultaneously.

A difficult problem while dealing with gestalts is the conjoined effect of two or more Gestalt principles operating at the same time on the same site. Modeling gestalts in such cases is especially challenging due to the complexity and ambiguity of the scene. Recently, attempts to discover how grouping principles interact were made in psychology and computer vision [Desolneux et al. 2002; Feldman 2003; Cao et al. 2007; Kubovy and van den Berg 2008]. These works model limited gestalt interactions, by finding computational means which are physiologically plausible. Kubovy and van den Berg [2008] explore the quantification of perceptual groupings formed conjointly by two grouping principles: similarity and proximity. Nevertheless, providing general computational means for modeling the interaction of multiple Gestalt principles remains a difficult challenge.

In this paper, we take a first step in developing a computational model for conjoining Gestalt rules. We model a subset of Gestalt rules and their mutual interaction for abstracting architectural line drawings. We choose to focus on architectural drawings since typically their visual elements are of rather low complexity and their spatial arrangement is strongly biased to the main axes (due to engineering considerations). Hence, architectural drawings consist of prevalent similarities, proximities and regularities among their el-
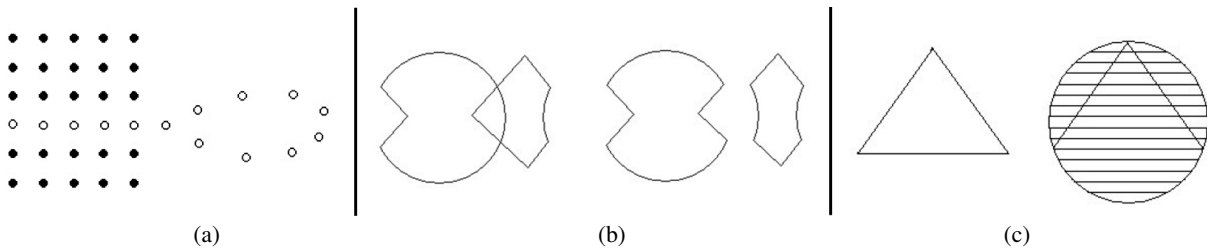
Figure 2: *Conjoining gestalts (from Kanizsa [1980]). Overlapping (a): white dots are elements of the grid (regularity) and simultaneously belong to a curve (continuity). Conflicting (b): continuity principle of two closed curves (b-left) conflicts with the symmetry principle (b-right). Masking (c): the basis of the triangle becomes invisible as it is embedded in a group of regularly parallel lines.*

ements, forming complex grouping configurations that can be described by Gestalt principles and their interaction. These rules define various perceptual scene groupings, naturally lending themselves to simplification and abstraction of architectural drawings.

Interacting gestalts rules cast a difficult problem since they operate simultaneously on common sites. They can compete, conflict, overlap, and mask with each other yielding complex visual phenomena. Up to now, psychologists are still studying how exactly multiple Gestalt principles interact. It is not in our scope to study the human visual system under these complex phenomena. Our goal is to quantitatively model and apply conjoining rules with respect to a specific type of data. Specifically, we propose a computational framework to facilitate the integration of five principal Gestalt rules, namely *similarity*, *proximity*, *continuity*, *closure*, and *regularity*, and then apply it to the abstraction of architectural line drawings. Due to the grouping nature of Gestalt rules, we formulate abstraction as a grouping optimization problem. The additive nature of Gestalt principles interactions (i.e., gestalts combine additively as showed in [Elder and Goldberg 2002; Feldman 2003; Kubovy and van den Berg 2008]) is modeled by a multi-label normalized graph cut formulation.

We apply conjoining Gestalt grouping rules, inspired by the computational model of Kubovy and van den Berg [2008]. We develop a global energy function that relates grouping strength, number of disjoint groups and inter-part characteristics to Gestalt principles. The minimal energy defines a global segmentation into disjoint groups while accounting for metrics that measure grouping in terms of Gestalt rules. The set of elements grouped is then abstracted by means of replacing them with simplified or abstract representatives. Our method computes a progressive series of abstractions from the finest, most detailed drawing, up to the coarsest most abstracted one (see Figure 1).

Our framework naturally lends itself to efficient abstraction of architectural line drawings conforming to perceptual grouping concepts. We demonstrate that our method succeeds in generating abstractions of architectural drawings which mimic the work of skillful artists. We also extend our method and show some basic results on mosaics and 3D buildings which are other families of objects that consist of discrete elements. The effectiveness of our method is evaluated by representing drawings in a range of scales, comparisons with previous methods and a user study.

2 Related Work

Perceptual Gestalt Grouping. Gestalt psychology is a theory on how humans perceive forms (figures or objects) instead of a collection of simple lines and curves. Wertheimer [1923] establishes the well-known Gestalt principles that describe strategies how human vision groups objects into forms. While early psychological stud-

ies focused mainly on qualitative and empirical studies, more recent psychological works study quantitative aspects of Gestalt principles and their conjoining. Desolneux *et al.* [2002] and Cao *et al.* [2007] formulate probabilistic quantities for distinctive Gestalt rules and utilize them for detecting collinearity, regularity and proximity in images. In his work, Feldman [2003] suggests a hierarchical representation for modeling proximity and collinearity grouping principles. He applies a minimal model theory for selecting the best grouping interpretation, preferring a maximally explained group. Kubovy and van den Berg [2008] present a probabilistic model for measuring proximity and similarity grouping quantities. Similar to us, they analyze and model the conjoint effect of grouping principles operating simultaneously. Nevertheless, their work focuses on interaction of only two principles performing on simple dot lattice configurations, while our goal is to provide a generic framework for resolving conjoining gestalts in the context of architectural abstraction. Claessens and Wagemans [2008] present a Bayesian model for contour detection using proximity and collinearity grouping. In contrast to ours, their method handles simplified interactions between two gestalts that perform as independent variables on dot lattices. In recent work by Cole *et al.* [2008; 2009], the authors study the correlation between shape conveying hand drawn lines and specific 3D shape features. Their work quantifies the strength of different line drawings at depicting shape.

Abstraction and Simplification of 2D Content. Gestalt effect is highly related to abstraction and simplification. By understanding how a collection of details aggregates, we can replace this collection with a simpler, coarser or abstract representation, while preserving the original semantics. In computer graphics and vision, Gestalt principles have been studied in the context of simplification and abstraction in several works. We focus here on works which address simplification using perceptual principles.

In an early work, DeCarlo and Santella [2002] compute image abstraction by preserving meaningful structures using an eye tracker to assist their image analysis. In their work, Mi *et al.* [2009] compute a decomposition of 2D shapes into parts using shape symmetries. Abstraction is achieved by removing parts according to their size.

A large amount of work exists on abstraction of line drawings. In the context of perceptual abstraction, Grabli *et al.* [2004] simplify line drawings using a complexity measure which accounts for stroke density and regularity variations. Nevertheless, the interrelations between density and regularity is not addressed. Barla *et al.* [2005; 2006], present algorithms for line drawing simplification and synthesis based on perceptual line grouping, accounting for proximity, color and continuation principles. Nevertheless, line stroke pairs are clustered greedily by selecting line pairs that satisfy one or more principles. Shesh and Chen [2008] present an efficient proximity measure for dynamic line grouping and simplification.

Perceptual grouping for synthesis purposes has been investigated recently. Ijiri *et al.* [2008] study element arrangement patterns by analyzing the relations between neighboring elements for the purpose of texture synthesis. Hurtut *et al.* [2009] perform analysis of element appearance for synthesizing 2D arrangements of stroke-based vectors. In their work, they measure the appearance of element qualitatively using a statistical model for finding meaningful appearance features.

Besides abstraction of 2D image content, one can also perform abstraction and simplification in other domains. Wang *et al.* [2004] perform abstraction of video sequences by semi-automatic segmentation of semantical contiguous volumes. In a recent work Mehra *et al.* [2009] create envelope shapes for complex 3D objects to guide their simplification.

Although these works address perceptual principles for simplification, they do not address the complex interrelations emerging from conjoining Gestalt principles using a computational model.

Abstraction of Architectural Drawings. Our motivation of utilizing conjoining Gestalt rules is for simplification of architectural drawings. Several works deal with abstraction and simplification of buildings and urban scenes for improving, clarifying and emphasizing the urban data. They play an important role in many location-based services, navigation and map generation applications (e.g., tourist maps).

Sidiropoulos and Vasilakos [2006] explore visualization methods for digital city representations. In their work they discuss various symbolic and realistic representations for urban visualization. Grabler *et al.* [2008] simplifies building appearance to de-emphasize less important buildings and reduce the complexity of tourist maps. Building complexity is measured using rectangularity and normals variation while simplification is performed using proximity of facets. Adabala *et al.* [2007; 2009] compute stylized maps and abstractions by straightening edges and modeling a facade detail variation (i.e., windows) using a combination of periodic "facade waveforms". Similarly, Loya *et al.* [2008] compute periodic features of building facades using Fourier series and render a reduced pattern to obtain simplification. In our work we utilize similar periodic waveform for modeling regular groups in the input. Finally, to improve perception of complex city areas, Glander *et al.* [2008; 2009] present a hierarchical abstraction in which buildings and streets are merged and removed by their proximity and size.

Although the above works consider implicitly, perceptual principles for architecture abstraction, our work is the first attempt to explicitly employ conjoining Gestalt rules for architecture abstraction. In the next section we describe our method in detail. Section 3 defines the Gestalt principles and their conjoining interactions. In section 4, we define quantitative measures for grouping elements to one or more gestalts. Section 5 describes our graph-cut formulation for modeling conjoining gestalts and Section 6 presents our simplification procedures for abstraction of Gestalt groups. Finally, in Section 7, we present our results, perform a user study and conclude.

3 Gestalt Basics and Interrelations

Gestalt principles describe how humans recognize a group of fine elements as a larger aggregate entity. This suggests that we can formulate the Gestalt phenomenon as a grouping problem and solve for an optimal grouping. We choose to model in this work a subset of Gestalt principles common in architectural drawings: similarity, proximity, continuity, closure, and regularity. These grouping principles are defined qualitatively in psychology as follows:
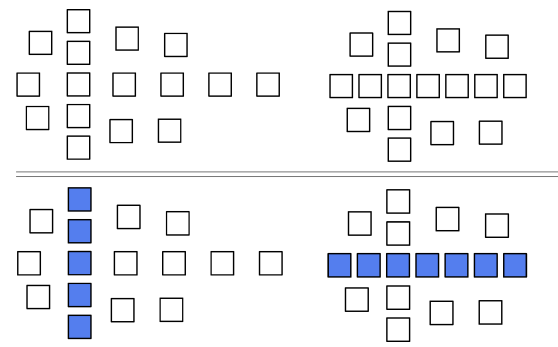
Figure 3: *Conflicting case 1: vertical (left) vs. horizontal (right) regularity gestalts. In top-left subfigure, vertical regularity is* masking *the horizontal one due to higher density. In top-right subfigure, horizontal regularity overrules the vertical due to an equal density although a larger group. Bottom row shows our grouping results(blue).*
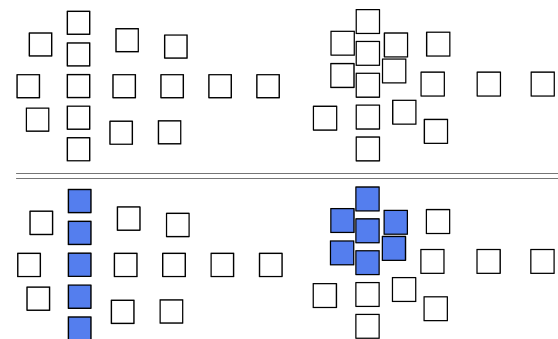

Figure 4: *Conflicting case 2: vertical regularity (left) vs. proximity (right) gestalts. In top-left subfigure, vertical regularity is* masking *proximity due to its strong regularity. In top-right, proximity overrules vertical regularity due to the stronger proximity. Bottom row shows our grouping results (blue).*
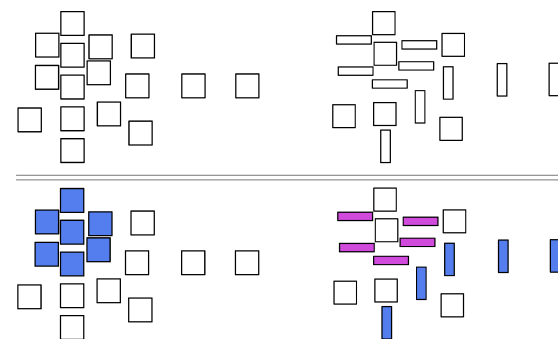
Figure 5: *Conflicting case 3: proximity (left) vs. shape similarity (right) gestalts. Top-right, shape similarity is* masking *proximity due to two groups of very similar shapes. Bottom row shows our grouping results (blue).*

- Similarity - parts which share visual characteristics such as shape, size or orientation can form a perceptive group.

- Proximity - parts which are closer together can be regarded as one group.

- Regularity - parts which are regularly spaced are seen as belonging together.
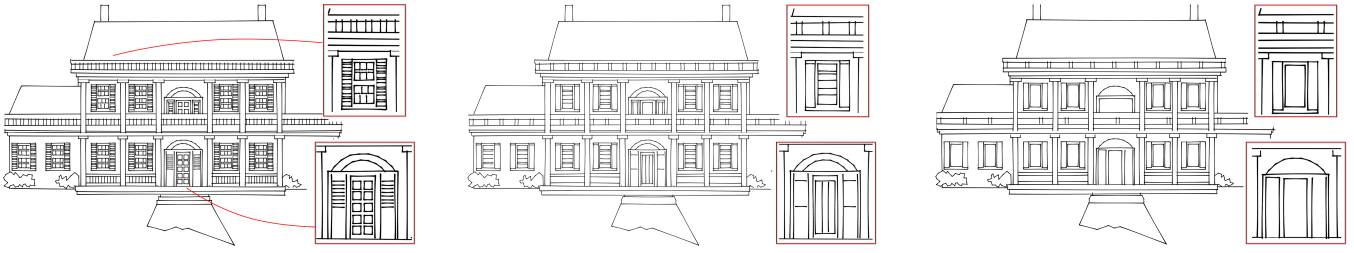
Figure 6: *Progressive abstraction of a complex facade based on conjoining gestalts. Zoom-ins of two different regions demonstrate our preservation of meaningful structures.*

- Continuity - preference for continuous shapes, thus seeing aligned disjoint elements as one group.

- Closure - if enough shape is indicated, the whole is obtained by filling missing data, thus closing simple figures.

Gestalt principles are stated as independent grouping rules as they start from the same building elements. When interactions between grouping rules occur (denoted as conjoining gestalts), the same scene might have different interpretations, which can lead to perception of sometimes incompatible groups in a given figure. The challenging phenomena of conjoining Gestalt principles were studied in the seminal work of Kanizsa [1980], where conjoining gestalts are described as gestalt principles in an equilibrium, struggling to give the final figure its organization (see Figure 2). The three cases psychology mentions are:

1. *Overlapping*: Two grouping principles act simultaneously on the same elements and give rise to two overlapping groups.

2. *Conflicting*: Both grouping principles are potentially active, but groups cannot exist simultaneously. Therefore, none of the grouping principles wins clearly leading to ambiguity as viewers can see both groupings.

3. *Masking*: Two conflicting grouping pricniples compete and one of them wins. The other one is inhibited.

There are many empirical studies on Gestalt behavior, but not many quantitative ones exist. Quantitative study on the interaction among multiple Gestalt principles is even scarce [Kubovy and van den Berg 2008]. Note that different rules may act at different levels and may interfere with each other. In our work, we quantify the affinity of an element to a gestalt according to quantitative measures. The conjoining gestalts interactions are modeled by formulating the spatial relationship among elements as a graph. Our computational model accounts for conflicting and masking interactions and resolves them by finding an optimal consistent grouping using a weighted energy minimization scheme. We currently do not handle overlapping since computationally it is difficult to differentiate from masking and conflicting phenomena.

Figures 3 – 5 demonstrate six scenarios in which one Gestalt rule masks another. Readers are encouraged to hide the bottom rows, and look at the top rows for a while to observe the forming of gestalts. In each input configuration (two for each case), there are at least two potential gestalts that can be observed, but at the same time, they compete with each other and the gestalt with stronger affinity masks the other. We pair into three cases, to accentuate *turning-points*, where a small change in the configuration results in a completely different gestalt. The illustrations aim to emphasize that the decisions taken can be quite complex, involving proximity, regularity and similarity. Our graph-cut solution resolves conflicts (colored bottom rows in Figures 3-5) and mimics the gestalts as

perceived by human. In Figure 7 we show a similar scenario of conflicting gestalts in a detailed window drawing.

4 Quantifying Gestalt Principles

Our input consists of 2D vector drawings of architectural scenes. We define *shape-elements* as closed connected polylines (open polylines for mosaics examples) which we automatically detect in the drawing. We first compute the spatial relationship among elements, so as to construct an associated graph. An element corresponds to a node in the graph while edges correspond to the spatial relationship among elements. We formulate Gestalt principles as probability functions and compute the probabilities of each element belonging to one or more gestalts.

The gestalts (hence groups) are the labels that are assigned to the nodes. We employ a multi-label graph cut (in Section 5) to partition this highly connected graph structure into optimal and consistent gestalts (groups of elements). Once the set of Gestalt-based groups is computed, the input detail drawing is simplified by abstracting/simplifying elements in the same gestalt and render them in various styles (Section 6).

Proximity Graph Structure. Our input is a vector drawing P consisting of polylines denoted p. Although they can be disjoint, intersect, or include each other (see Figure 6), we assume that a polyline represents a unique shape element. We first compute a proximity graph G_p, with each element corresponding to a node in the graph. For each element p_i, we find its k-closest neighbors in the drawing and connect the corresponding nodes in the graph with edges e_{ij}, associated with the Hausdorff-distance between p_i and p_j, defined as:
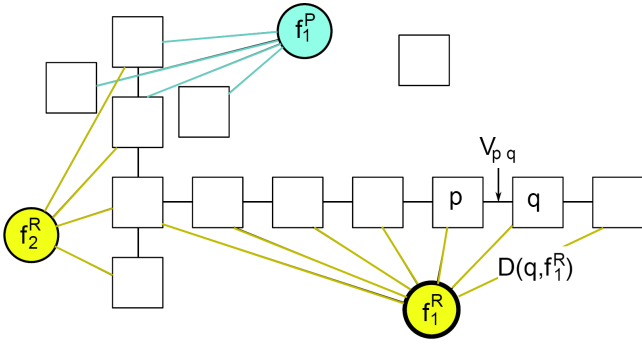
$$d(p_i, p_j) = \max\{d_H(p_i, p_j), d_H(p_j, p_i)\},$$

and

$$d_H(p_i, p_j) = \max_{\forall v_i \in p_i} \{ \min_{v_j \in p_j} \{\|v_i - v_j\|_2\}\},$$
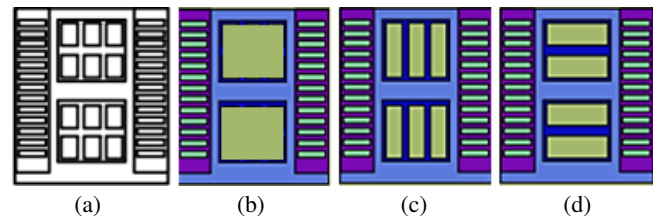
Figure 7: *In a window (a) from Figure 6 conjoining gestalts compete by: similarity (b), vertical regularity (c) and horizontal regularity (d) which is the winning gestalt.*

Figure 8: *A simple illustration of our graph construction. An element p is connected to its neighbor q by edges defining the smoothness term $V_{p,q}$. We compute potential gestalts denoted here by f_1^R, f_2^R, f_1^P where f_*^R denotes regularity type gestalt and f_*^P, proximity. An element is assigned to potential gestalts using a data cost $D(q,f)$, measuring the penalty of assigning label f to element q.*

where v_i and v_j are vertices in p_i and p_j, respectively. This distance is assigned as the weight of the edge. Figure 8 shows a simple example of our graph construction.

Quantifying Affinity. Initially, we loosely define *potential groupings* based on Gestalt rules, which will serve as graph labels in our graph-cut formulation. For each potential group, we quantify the affinity of an element p_i to it by computing a data term that measures the probability of p_i to belong to a gestalt (i.e. being assigned with that label). Additionally, we prioritize labels and assign each label a cost by its grouping strength measuring the affinity in metrics we define below. Next, we define the potential groups.

Proximity gestalts are computed by detecting groups of connected elements in G_p with edge lengths below a threshold t_p. Thus, a proximity group is defined as:

$$L_P = \bigcup\{p_i, p_j\} \mid d(p_i, p_j) < t_p.$$

We observe that elements may intersect or enclose each other (for example, a small window inside a door). Since proximity relations are undefined in such cases, we avoid considering such elements in one proximity gestalt. I.e., for enclosed elements, we do not consider proximity outside their enclosing element. We enforce this by automatically detecting and storing intersection and inclusion relations between shape elements in the drawing.

Similarity gestalts are computed by detecting groups which share a high shape similarity. Since we mainly focus on architectural elements, it was sufficient to measure similarity by comparing the aspect ratio of bounding boxes of elements. If required, complex shape similarity metrics such as the transformation-invariant shape context [Belongie et al. 2002] can be employed. A similarity group is defined as:

$$L_S = \bigcup\{p_i, p_j\} \mid \frac{R(H_i, H_j) + R(W_i, W_j)}{2} > t_s,$$

where W_i and H_i are width and heights of element p_i; and

$$R(a,b) = \begin{cases} a/b & \text{if } a < b, \\ b/a & \text{otherwise.} \end{cases}$$
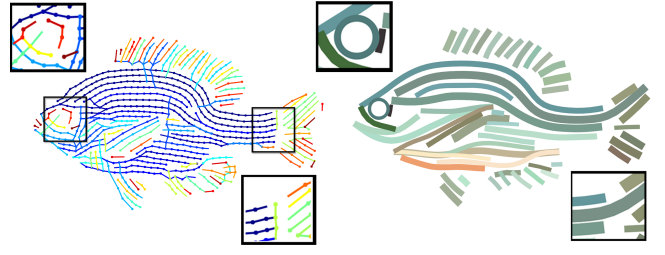
Figure 9: *An example of closure and continuity gestalts. RANSAC detects a circle that is fitted to the fish eye and defines a strong closure gestalt. Similarly, continuity gestalts occur at the tail defining continuous lines.*

Regularity gestalts are computed by detecting regular structures in the scene. We define the regularity as a group of elements (larger than 2) which are positioned at regular intervals along a certain direction. Although sophisticated symmetry analysis [Liu et al. 2004] can be employed, we reduce our search space to 1D regularities along X and Y axes following the Manhattan-world assumption, commonly applied to architectural models. We perform a 2D frequency decomposition of the input scene into axis aligned vertical and horizontal dominant frequencies similar to [Adabala 2009]. Specifically, we subdivide the 2D image into a set of horizontal and vertical non-overlapping tiles and compute the 1D FFT for each tile. We filter low magnitude frequencies thus obtaining a set of dominant vertical and horizontal regular candidates corresponding to high magnitude frequencies. We define regularity potential gestalts as:

$$L_R = \bigcup\{p_i\} \mid \xi(p_i) > t_r,$$

where $\xi(\text{p}_i)$ is a detected dominant horizontal or vertical frequency magnitude of element p_i.

Closure gestalts refers to a group of elements that forms a simple shape. *Continuity* is a special case in which elements lie on a line or curve. We compute closure potential groups by fitting simple geometric primitives θ, such as straight lines (for continuity), circles and squares (for closure) to the scene elements. For each primitive type, we fit it to drawing elements using RANSAC and group together elements with a high fitting score as:

$$L_C = \bigcup\{p_i\} \mid \text{fit}(p_i, \theta) > t_c,$$

where $\text{fit}()$ measures the fitting quality of p_i to θ by counting the number of points (i.e. polyline vertices) which are within an ε distance from θ. See Figure 9 for an example of closure gestalts, by fitting a circle (top zoom) and straight lines (bottom zoom) to polyline vertices. The resulting fitted vertices are grouped together and simplified as will be described in section 6.

With the above definitions, we form many potential groups (gestalts) from the input drawing, each corresponding to a distinct label. Obviously, an element may belong to several gestalts (labels) even by the same Gestalt rule (in Figure 8 labels f_1^R, f_2^R). Thus, the scene is over-segmented into groups which possibly interact. We use graph-cut to resolve interactions and achieve a consistent segmentation of the scene into groups with minimal energy. During our graph cut optimization, potential groupings may break into separate parts due to conflicts.
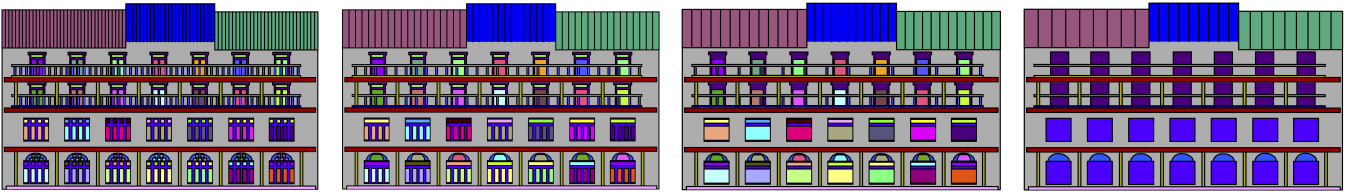
Figure 10: *A sequence of abstraction steps. We color-code corresponding element groupings to visualize the computed gestalts. Two abstraction operations are performed, (a) summarization by reducing railings number in fences, and (b) embracing by replacing window elements with enclosing object. Although railings and doors overlap, their interaction is solved as railings are grouped together by regularity gestalt.*

5 Conjoining Gestalts via Graph Cut

Resolving conjoining gestalts in a scene is equivalent to finding a consistent segmentation of elements into groups which comply with Gestalt rules. We formulate the problem as a multi-label normalized graph cut minimization. As an element p can potentially belong to many gestalts, it gets assigned data cost for different gestalt labels (colored nodes in Figure 8). Given n elements, k labels and $n \cdot k$ costs, finding the minimal assignment is a combinatorial problem and typically NP-hard. Instead, we follow Delong *et al.* [2010] and use an approximate multi-label graph-cut energy minimization.

We compute an assignment of labels f_p to elements $p \in P$ such that the joint labeling f minimizes an objective function $E(f)$. Our function consists of three terms: data, smoothness, and label costs. Kubovy and van den Berg [2008] showed that proximity grouping strength decays exponentially with Euclidean distance. We explored several metrics for quantifying gestalts and have found them to behave very similar in presence of conjoining gestalts. We select simplified gestalt metrics which improve our energy minimization convergence. Without loss of generality, we assume all label costs to be normalized and bounded.

Label Cost. Label cost penalizes overly-complex models and favors the explanation of the input scene with fewest and cheapest labels. The label cost function is defined as:

$$F_{\text{cost}} = \sum_{l \in L} h_l \cdot \delta_l(f)$$

with L being the set of labels, h_l a non-negative label cost of label l and $\delta_l()$ an indicator function:

$$\delta_l(f) = \begin{cases} 1 & \exists p : f_p = l, \\ 0 & \text{otherwise} \end{cases}$$

In the following, we define h_l as the label cost measuring the gestalt affinity for each specific Gestalt rule (normalized by mapping min-max to $0 - 1$ respectively):

- For proximity gestalts, label cost is measured by the inverse density defined as the area difference between the union of shapes and their convex hull: $h_l =_{p \in L_P} CH(p) - \bigcup(p)$

- For similarity gestalts, label cost is defined by shape similarity variance against an arbitrary shape within the gestalt: $h_l =_{(p_i, p_j) \in L_S} \text{var}(R(H_i, H_j) + R(W_i, W_j))$

- Regularity label cost is measured as the inverse density multiplied by the elements distance variance from the perfect frequency pattern and inverse number of elements: $h_l =_{p \in L_R} (CH(p) - \bigcup(p)) \times \text{var}(p, \xi) \times \|L_R\|^{-1}$

- Continuity and closure label cost is measured by the fitting quality, defined as the distance variance of group members to fitted geometric primitive: $h_l =_{p \in L_C} \text{var}(\text{fit}(p_i, \theta))$

Smoothness Cost. Smoothness term measures the spatial correlation of neighboring elements. Elements with a smaller distance have a much higher probability to belong to the same gestalt than those distant ones. This is defined in our energy minimization scheme as the smoothness term. Between two neighboring elements p and q, the smoothness energy term is defined by the inverse Euclidean Hausdorff-distance between p and q (normalized by mapping min-max to $0 - 1$ respectively):

$$V_{pq} = d(p, q)^{-1}$$

Data Cost. Data cost $D(p, f_p)$ measures how well an element p fits to a gestalt f_p, (normalized by mapping min-max to $0 - 1$). Although theoretically each element and potential gestalt define a data cost, we take only elements within a threshold from each gestalt. This does not affect the solution since elements too distant from a gestalt will not be grouped together. We define the data cost for each Gestalt type as follows:

Proximity data cost is simply defined as the closest distance of the element p from the proximity group L_P defined as:

$$D(p, f_p) = \min_{q \in L_P} d(p, q)$$

Similarity data cost is defined as the average shape similarity distance of p to elements in the similarity group L_S,

$$D(p, f_p) = \frac{1}{|L_S|} \sum_{q \in L_S} \{R(H_p, H_q) + R(W_p, W_q)\}$$

Regularity data cost is defined as the distance from the regular pattern defined by L_R. Given a regular pattern, we compute the distance $d(p, p')$ of p from the ideal element p' that perfectly aligns with the regular pattern.

$$D(p, f_p) = d(p, p')$$

Continuity (closure) data cost is measured as the closest distance of an element to the fitted geometric primitive. Given a continuity group L_C which defines a fitted geometric primitive, $d'(p, *)$ is the closest distance of p to the primitive defined by L_C.

$$D(p, f_p) = d'(p, *)$$

Optimization. Hence, the overall energy function is:

$$E(f) = \sum_{p \in P} D(p, f) + \sum_{p,q \in N} V_{p,q} + \sum_{l \in L} h_l \cdot \delta_l(f)$$

Finding a solution to this labeling problem is optimized using a multi-label normalized graph-cut algorithm as proposed by Delong et al. [2010]. Theoretically, we can compute data costs $D(p, f)$ for each p and l, using the complete set of possible groupings. However, this would be too large and instead use the thresholding of the formed gestalts (t_p, t_s, t_r) to limit data cost computation to only group members.

Iterative Conflict Resolution. As mentioned in Section 3, conflicts in a scene occur when partial gestalts are hidden by other partial gestalts and give equivalent or better explanation of the scene. The natural outcome of our graph-cut minimization is a segmentation into minimal cost gestalt assignment. Thus, when gestalts compete due to conflicts, naturally the strongest gestalt in terms of its energy terms wins. Thus, the model clearly complies with the conflicting and masking phenomena.

Still, our graph-cut assigns labels to overruled groups resulting in partial gestalts. We detect such assignments and remove them. To do so, we evaluate the graph-cut assignment by measuring the total data cost term of each group before and after graph-cut. If there is a large drop in data cost, it means the group has been overruled by another gestalt and we denote it as invalid and mark its elements as unassigned. Note that overruled gestalts may still create new (smaller) gestalts by themselves and therefore we repeat graph-cut computation iteratively. In each iteration we compute potential groups, perform graph cut and detect valid gestalts. Next we remove the valid groups from the graph and repeat the whole process on the remaining elements. We stop when no new gestalts can be formed.

In conflict cases, where two gestalts equally compete on common elements, graph cut will choose one rule arbitrarily. The phenomena of both groups breaking together seldom occurs since graph-cut minimizes the number of distinct labels and data cost, thus preferring large low cost groups.

Figures 3- 5 shows conflicting element configurations (upper row) and our optimization results (lower row). In Figure 3(top), two candidate vertical and horizontal regularity gestalts compete with each other as they share a common element. In the left example, the vertical gestalt wins as its regularity is denser thus having lower label cost. On the right, the horizontal gestalt wins since both vertical and horizontal gestalts have equal density but the horizontal gestalt has a higher element number thus a lower label cost. The colored elements in the low rows of the figure show our computed gestalt. In Figure 4(right), elements congregate together, leading to competition between rules of (vertical) regularity and proximity. Our method selects the proximity gestalt (lower right), due to the superiority of proximity. If we change the shapes of elements as in Figure 5(right), the previously computed gestalt in Figure 5(lower left) is split due to the superiority of similarity.

6 Visual Abstraction

From the computed gestalts, we can apply different abstraction methods to achieve different styles. Unlike previous abstraction techniques in non-photorealistic rendering, our gestalt-based approach *decouples* the identification of abstracting regions from abstraction styles. In other words, the same gestalt can be presented
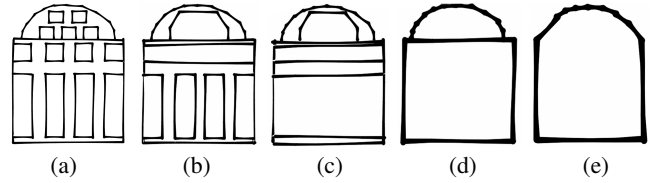
Figure 11: *Gestalts are abstracted by creating a simple representation for the overall form. Here we show progressive grouping (from left to right) by regularity, proximity and similarity. Since abstraction is governed by an LOD threshold, we first simplify smaller gestalts, and later larger ones.*

with different abstraction styles. We define abstraction styles in a content dependent manner.

Abstracting Architectural Drawing. We propose two types of operators, *embracing* and *summarization*, for abstracting architectural drawing. Embracing replaces elements in a gestalt by a simplified enclosing shape. Summarization represents the repeated elements in a gestalt with a smaller number of repeated elements. Our choice of abstraction operator depends on the number of repeated elements in a gestalt.

If a gestalt contains more than t_k repeated elements (in our implementation $t_k = 20$), we apply a summarization operator which gradually reduces the number of elements. Otherwise gestalts are abstracted using an embracing operation. We use the convex hull of the elements in a gestalt for its embracing. We observe that architectural drawings typically consist of dominant horizontal and vertical directions denoted as Manhattan World alignment. Therefore, if an enclosing convex hull edge is within a threshold of 10 degrees with vertical/horizontal direction, we align it with that direction.

Figures 6, and 10 demonstrate progressive simplification sequences using the two different ways of abstraction. The fine details on the door and windows are abstracted using embracing, while the fences are abstracted with summarization operation.

Figure 11 shows a window from Figure 10 being progressively simplified, where convex hulls and axis aligned boxes are used for embracing. This interesting example demonstrates the effects of different Gestalt principles during abstraction. From (a) to (b), the rules of proximity and similarity take part in the upper part of the window, while the rules of regularity and similarity take effect in the lower part. The rule of regularity continues to take effect from (b) to (c). From (c) to (d), the rules of similarity and proximity take part, and finally in (e) the rule of proximity groups the half disk with the box. For abstraction, the embracing operator applies convex hulls (upper window) and boxes (lower window) for the elements.

Level-of-Detail To generate progressively simplified results, we can repeatedly apply the gestalt computation in a progressive manner. That is, we take the current computed gestalt simplification result as input elements, and apply the gestalt simplification again. This process repeats until reaching the desired level or no new gestalts arise.

Gestalts computation does not explicitly account for progressive level-of-details although locally, using our label cost formulation, denser groups are superior to sparser ones. Hence, we can obtain gestalts that are arbitrarily large. In order to achieve a level-of-detail hierarchy we define a threshold parameter $0 \le t_{\text{LOD}} \le 1$ and limit forming gestalts that contain neighbor elements p, q with distance $d(p, w) \ge t_{\text{LOD}} \cdot D_{bbox}$, where D_{bbox} is the drawing's bounding

box.

We can synchronize the level-of-detail with screen resolution by measuring the D_{bbox} in screen space coordinates denoted by D'_{bbox} and defining $t_{\text{LOD}} = \frac{2}{D'_{bbox}}$. Thus, our coarsest resolution is achieved when $D'_{bbox} = 2$ pixels. As t_{LOD} increases, larger gestalts are permitted to form and hence coarser level-of-details. Figures 1, 13 and 15 show progressively simplified examples in reducing scale. Note that the core structure is preserved and still clearly apparent even when the building is significantly scaled down.

Abstracting Mosaics In architectural drawings, the rules of proximity, similarity, and regularity dominate, while the rules of continuity and closure seldom apply. To demonstrate the effect of these rules, we extend our application to non-architectural content. In particular, we focus on mosaic art since artists cleverly position and orient the discrete tiles to exhibit the structure through continuation and closure effects. By computing gestalts, we can extract the structure, mostly due to rules of continuity and closure. Starting from a set of colored tiles given as polygonal elements, we cluster them by color similarity. In a cluster, we connect between centers of neighbor tiles yielding a set of disjoint polylines. Figure 9(left) shows our input polylines extracted from a fish mosaic (in Figure 16(left)).

We show the continuity and closure gestalts obtained at different level-of-details on the fish mosaic. Polylines representing different tile clusters get grouped together mostly by continuity and closure gestalts. Primitive shapes as circles and straight lines are fitted to the polyline elements yielding closure and continuity gestalts respectively. For abstraction purposes, we compute for each gestalt a representing polyline by averaging vertex positions and a width that encloses the gestalt elements. We draw the new lines with average color of elements inside the gestalt.

7 Results

We have applied our gestalt-based abstraction on a variety of architectural drawings, exploring the behavior of our model in presence of complex conjoining gestalts. In all our experiments we have used the following thresholds of potential gestalts (see 4): proximity $t_p = 10$ pixels, similarity $t_s = 0.8$, regularity frequency magnitude $t_r = 0.7$ and closure fitting $t_c = 5$ with $\varepsilon = 2$ pixels. The average computational time for a coarse-to-fine abstraction sequence is 3 minutes and maximal time is 5 minutes.

In Figure 12, we evaluate our method by comparing our result against several manual and automatic techniques on the Taj Mahal drawing. To normalize the comparison, we bounded all simplifications by the same amount of geometric detail as measured by the amount of present line geometry. In Figure 12(a) is the result of a professional hand-drawn abstraction from a website that teaches drawing abstract buildings (http://tlc.howstuffworks.com/family/how-to-draw-buildings6.htm). In Figure 12(b), we asked an amateur artist to draw a simplified version while maintaining important structures. Figure 12(c) is the result by the technique of Shesh and Chen [2008] and (d) shows the result of applying proximity based simplification of geometric elements. Finally our result is in (e). Both ours (e) and the expert artist simplification (a) preserve the major structures and are comparable in terms of the visual quality. Nevertheless, Shesh and Chen [2008] (c) and geometric simplification (d) use low-level distance metrics for abstraction and fail to preserve the important high-level structures as in ours. Figure 13 shows a detailed sequence of our abstraction of the Taj Mahal. Our method can handle

fine building details with curved structures and thus is not limited to grid-based rectangular architecture.

In order to measure the effectiveness of our method we have run a thorough user study. In this study we have asked a group of 200 users to view simplified drawings and choose the one that best represents the original input. For that, we have built a test case containing twenty architectural drawings simplified by our method, hand-drawn by an artist and a straightforward geometric simplification technique (in Figures 12 (b), (d) and (e)). We compare equivalent simplifications, by measuring the amount of geometric detail and choosing drawings with similar amount of detail. As expected, users preferred our simplification over the straight-forward geometric approach with 74% votings for us against 26% for the geometric approach. We were surprised to discover that our method has done better even against the artist with 66% votings for us against 34% for the artist.

Since our method simplifies architectural drawings while preserving meaningful structures, we demonstrate in Figure 14 an immediate application of our method for computing thumbnail directories with increased visual perception. Figure 15 shows a building being zoomed out from left to right. Note how the scaled-down simplified versions preserve the major features and structures, and hence, can serve as representative *icons* of the original buildings. In Figures 1, 6, 11, 13 and 15 we render results using an NPR sketchy style for visual emphasis of abstraction. We include all results using a clean rendering style in the supplemental material.

Finally, in the mosaic example (Figure 16), the continuity/closure gestalts play a dominant role in the simplification sequence (see also Figure 9). Here, a circle is fitted to the fish eye during the closure gestalt computation. The strength of this gestalt preserves the fitted circle even after several iterations of simplification.

8 Concluding Remarks

In this paper, we have proposed a framework that models several prominent Gestalt principles and competition/conflicts among them. The framework is formed in a computational model that is realized with graph cuts. The effectiveness of this framework has been demonstrated in summarizing and abstracting architectural elements, with extended applicability to other objects, illustrated on mosaics arts.

The presented work represents still a first attempt to modeling the complex interaction among multiple Gestalt principles. We will continue to benefit findings from the psychological domain. The proposed computational framework can be easily adapted to confirm to psychological findings and extended to model other Gestalt behaviors. We believe that tools have be built for this framework and can be used to facilitate psychological experiments and leading to new findings.

While in this paper we demonstrate how our Gestalt analysis benefits abstraction for the computer graphics purpose, naturally the proposed method can also be effective for scene understanding in general, and further for bettering visual communication.
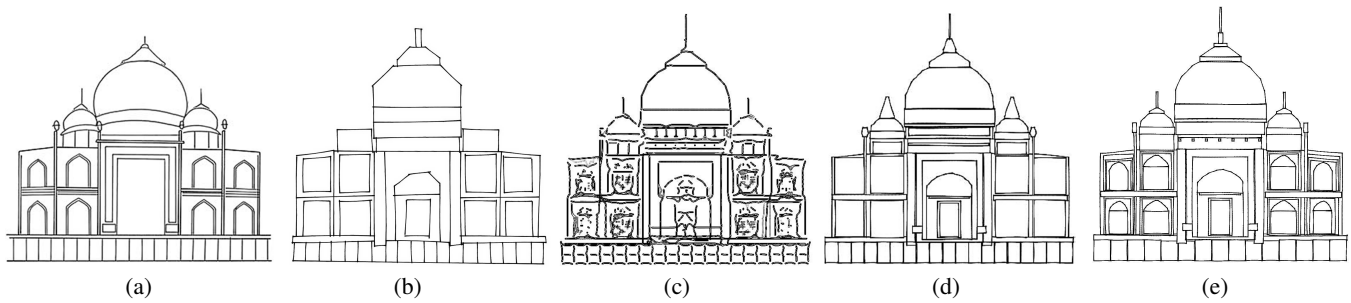
Figure 12: *Comparing various abstraction techniques on the Taj Mahal (see Figure 13) . (a) shows a professional hand-drawn abstraction, (b) is a hand-drawn abstraction by an amateur artist, (c) shows state-of-the-art simplification using Shesh and Chen [2008], (d) is geometry simplification by proximity and (e) shows our gestalt-based abstraction result.*


Figure 13: *A sequence of gestalt based abstractions on the highly-detailed Taj Mahal.*
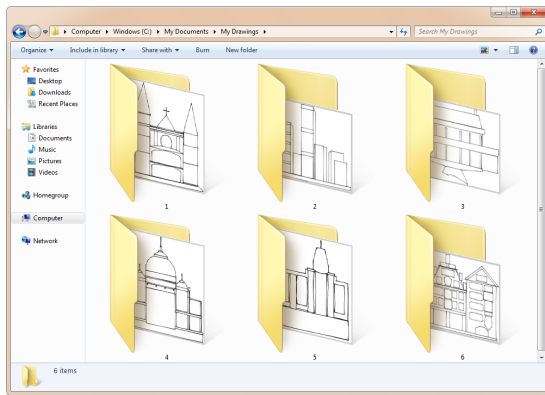
Figure 14: *An architecture directory is illustrated using thumbnails of simplified drawings from our method.*

References

ADABALA, N., VARMA, M., TOYAMA, K., AND BANGALORE, S. N. 2007. Computer aided generation of stylized maps. *Comp. Anim. Virtual Worlds 18*, 133–140.

ADABALA, N. 2009. Building representation in oblique-view maps of modern urban areas. *Cartographic Journal 46*, 2, 104–114.

BARLA, P., THOLLOT, J., AND SILLION, F. X. 2005. Geometric clustering for line drawing simplification. In *In Proceedings of the Eurographics Symposium on Rendering*.

BARLA, P., BRESLAV, S., THOLLOT, J., SILLION, F. X., AND MARKOSIAN, L. 2006. Stroke pattern analysis and synthesis. *Comp. Graph. Forum 25*, 3, 663–671.

BELONGIE, S., MALIK, J., AND PUZICHA, J. 2002. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 509–522.

CAO, F., DELON, J., DESOLNEUX, A., MUSE, P., AND SUR, F. 2007. A unified framework for detecting groups and application to shape recognition. *J Math Imaging Vis 27*, 91–119.

CLAESSENS, P., AND WAGEMANS, J. 2008. A bayesian framework for cue integration in multistable grouping: Proximity, collinearity, and orientation priors in zigzag lattices. *Journal of Vision 8*, 4, 33:1–23.

COLE, F., GOLOVINSKIY, A., LIMPAECHER, A., BARROS, H., FINKELSTEIN, A., FUNKHOUSER, T., AND RUSINKIEWICZ, S. 2008. Where do people draw lines? In *ACM SIGGRAPH 2008 papers*, ACM, 1–11.

COLE, F., SANIK, K., DECARLO, D., FINKELSTEIN, A., FUNKHOUSER, T., RUSINKIEWICZ, S., AND SINGH, M. 2009. How well do line drawings depict shape? In *ACM Transactions on Graphics (TOG)*, vol. 28, ACM, 28.

DECARLO, D., AND SANTELLA, A. 2002. Stylization and abstraction of photographs. In *ACM Transactions on Graphics (TOG)*, vol. 21, ACM, 769–776.

DELONG, A., OSOKIN, A., ISACK, H. N., AND BOYKOV, Y. 2010. Fast approximate energy minimization with label costs. In *CVPR*, 2173–2180.
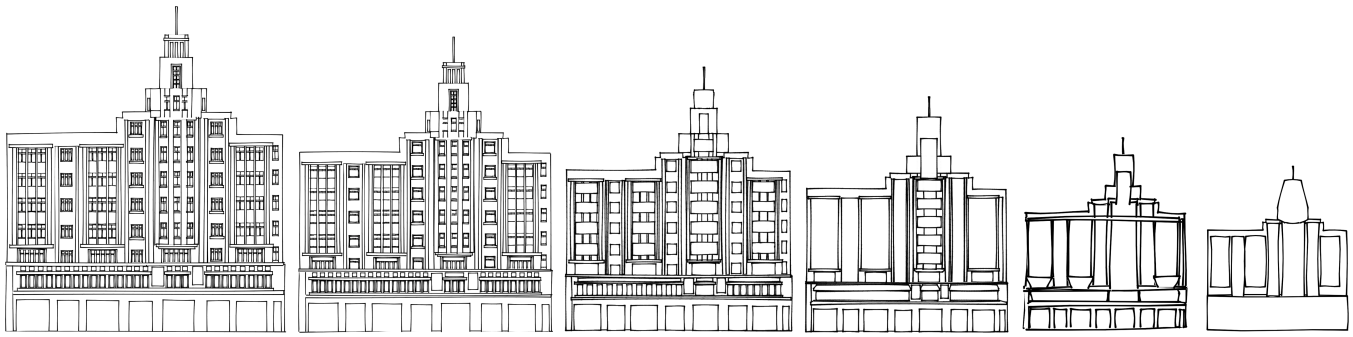
Figure 15: *Our method models the conjoining gestalts and correctly groups horizontal and vertical window structures and formations from a complex building facade.*

Figure 16: *Progressive gestalt abstraction of a fish mosaic. The major curve structures are identified and preserved through abstraction.*

DESOLNEUX, A., MOISAN, L., AND MICHEL MOREL, J. 2002. *Gestalt theory and Computer Vision*.

DRORI, I., COHEN-OR, D., AND YESHURUN, H. 2003. Fragment-based image completion. In *Proc. ACM SIGGRAPH 2003*, 303–312.

ELDER, J. H., AND GOLDBERG, R. M. 2002. Ecological statistics of gestalt laws for the perceptual organization of contours. *Journal of Vision*, 4.

FELDMAN, J. 2003. Perceptual grouping by selection of a logically minimal model. *International Journal of Computer Vision 55*, 1, 5–25.

GLANDER, T., AND DLLNER, J. 2009. Abstract representations for interactive visualization of virtual 3d city models. *Computers, Environment and Urban Systems 33*, 5, 375 – 387.

GLANDER, T., AND DÖLLNER, J. 2008. Techniques for generalizing building geometry of complex virtual 3d city models. In *Advances in 3D Geoinformation Systems*, Springer, P. van Oosterom, S. Zlatanova, F. Penninga, and E. M. Fendel, Eds., Lecture Notes in Geoinformation and Cartography, 381–400.

GRABLER, F., AGRAWALA, M., SUMNER, R. W., AND PAULY, M. 2008. Automatic generation of tourist maps. *ACM Trans. Graph. 27* (August), 100:1–100:11.

GRABLI, S., DURAND, F., AND SILLION, F. 2004. Density measure for line-drawing simplification. In *Proceedings of Pacific Graphics*.

HURTUT, T., LANDES, P., THOLLOT, J., GOUSSEAU, Y., DROUILLHET, R., AND COEURJOLLY, J. 2009. Appearance-guided synthesis of element arrangements by example. In *Proceedings of the 7th International Symposium on Non-Photorealistic Animation and Rendering*, ACM, 51–60.

IJIRI, T., MECH, R., IGARASHI, T., AND MILLER, G. 2008. An Example-based Procedural System for Element Arrangement. In *Computer Graphics Forum*, vol. 27, Wiley Online Library, 429–436.

KANIZSA, G. 1980. *Grammatica del Vedere*.

KUBOVY, M., AND VAN DEN BERG, M. 2008. The whole is equal to the sum of its parts: A probabilistic model of grouping by proximity and similarity in regular patterns. *Psychological Review 1*.

LIU, Y., COLLINS, R. T., AND TSIN, Y. 2004. A computational model for periodic pattern perception based on frieze and wallpaper groups. *IEEE Trans. Pattern Anal. Mach. Intell. 26*, 3, 354–371.

LOYA, A., ADABALA, N., DAS, A., AND MISHRA, P. 2008. A practical approach to image-guided building facade abstraction. In *Computer Graphics International 2008*.

MEHRA, R., ZHOU, Q., LONG, J., SHEFFER, A., GOOCH, A., AND MITRA, N. J. 2009. Abstraction of man-made shapes. *ACM Transactions on Graphics 28*, 5, 137:1–137:10.

MI, X., DECARLO, D., AND STONE, M. 2009. Abstraction of 2D shapes in terms of parts. In *Proceedings of the 7th International Symposium on Non-Photorealistic Animation and Rendering*, ACM, 15–24.

MITRA, N. J., CHU, H.-K., LEE, T.-Y., WOLF, L., YESHURUN, H., AND COHEN-OR, D. 2009. Emerging images. *ACM Transactions on Graphics 28*, 5 (December), 163:1–163:8.

SHESH, A., AND CHEN, B. 2008. Efficient and dynamic simplification of line drawings. *Comput. Graph. Forum 27*, 2, 537–545.

SIDIROPOULOS, G., AND VASILAKOS, A. 2006. Ultra-real or symbolic visualization? the case of the city through time. *Computers & Graphics 30*, 2, 299 – 310.

WANG, J., XU, Y., SHUM, H.-Y., AND COHEN, M. F. 2004. Video tooning. *ACM Trans. Graph. 23* (August), 574–583.

WERTHEIMER, M. 1923. Untersuchungen zur lehre der gestalt ii,. *Psychol. Forsch.*, 4, 301–350. Translation published as Laws of Organization in Perceptual Forms, in: W. Ellis, A Source Book of Gestalt Psychology, Routledge and Kegan Paul, London, 1938, pp. 71–88.