

Automatic Generation of Tourist Maps

Floraine Grabler Maneesh Agrawala
University of California, Berkeley

Robert W. Sumner Mark Pauly
Applied Geometry Group, ETH Zurich

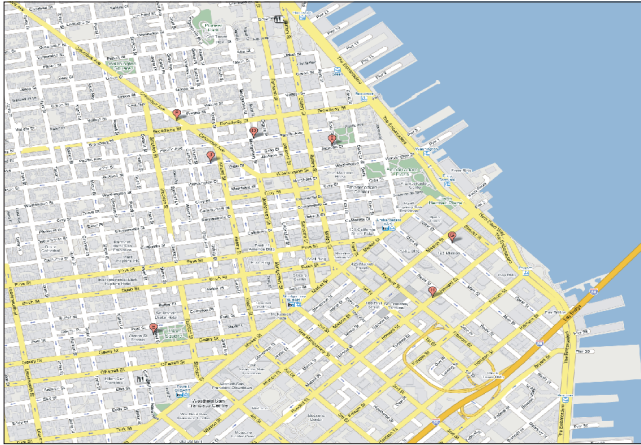


Figure 1: (left) A digital map of San Francisco from Microsoft Live (www.live.com) marks the location of a few user-chosen landmarks using pushpins. It is impossible for a tourist to recognize buildings without further visual information. (right) A tourist map generated by our system automatically includes contextual landmarks in addition to the category of buildings (i.e. restaurants, shopping, etc.) chosen by the user. Moreover the map depicts building facades so that tourists can more easily identify the buildings.

Abstract

Tourist maps are essential resources for visitors to an unfamiliar city because they visually highlight landmarks and other points of interest. Yet, hand-designed maps are static representations that cannot adapt to the needs and tastes of the individual tourist. In this paper we present an automated system for designing tourist maps that selects and highlights the information that is most important to tourists. Our system determines the salience of map elements using bottom-up vision-based image analysis and top-down web-based information extraction techniques. It then generates a map that emphasizes the most important elements, using a combination of multiperspective rendering to increase visibility of streets and landmarks, and cartographic generalization techniques such as simplification, deformation, and displacement to emphasize landmarks and de-emphasize less important buildings. We show a number of automatically generated tourist maps of San Francisco and compare them to existing automated and manual approaches.

Keywords: Visualization, Non-Realistic Rendering, WWW Applications

1 Introduction

Tourist maps are essential resources for visitors to an unfamiliar city because they visually highlight landmarks and other points of interest such as museums, restaurants, parks, and shopping districts. The most effective tourist maps are carefully designed to present this information so that visitors can easily navigate to the places they are

most interested in. Yet, designing and rendering such tourist maps is a time-consuming process that requires expert map design skills.

Moreover, hand-designed maps are static representations that cannot adapt to the needs and tastes of the individual tourist. Points of interest can differ significantly from person to person. While visually distinctive buildings and environmental features can serve as general-purpose landmarks, one tourist may be most interested in shopping, while another may want to see nearby restaurants. In designing a map, the first challenge for mapmakers is to determine the importance of these elements to the tourists that will use the map.

Tourists often use a combination of public transportation and walking to move from one place to another. Therefore, an effective tourist map must provide rich visual representations of landmarks and points of interest to help tourists quickly identify where they are located and determine the best (most interesting) route to their destination. For example, mapmakers use multi-perspective rendering to increase the visibility of streets and landmarks. Similarly, they use a variety of cartographic generalization techniques to increase the recognizability of landmarks and emphasize the most important elements in the map. Thus, the second challenge for mapmakers is to choose a set of rendering and cartographic generalization techniques that emphasize the most important landmarks and points of interest while de-emphasizing or eliminating irrelevant elements.

In the last decade, digital maps such as those provided by Microsoft Live (www.live.com), Google Maps (maps.google.com) and Yahoo Maps (maps.yahoo.com) have become increasingly popular. Cities are usually in a perpetual state of construction, renovation and renewal in which streets and buildings are created, destroyed and reshaped. One advantage of such digital maps over hand-designed maps is that they are based on continually updated geometric models of the streets and buildings in the city and therefore usually reflect the most up-to-date information. However, these digital maps do not adequately address the two primary challenges of creating tourist maps and suffer from several problems, including:

Poor recognizability of landmarks. Digital maps often leave out the buildings altogether or show only their footprints. While the location of user selected landmarks can be highlighted using pushpins, it is impossible for a user to recognize the landmark without

further visual information.

No context information. Digital maps do not provide context information because they do not highlight visually distinctive buildings that can help users orient themselves while navigating through the city.

Poor generalization. Digital maps do not generalize less important buildings of the city. All buildings are presented at the same level of detail, making it difficult for tourists to filter out the information they need. Thus, the maps appear overly complex.

In this paper we present an automated system for generating tourist maps that addresses the two primary mapmaking challenges highlighted earlier, and significantly reduces the problems inherent in current digital maps (see Figures 1 and 14). The input to our system consists of a geometric model of a city, including streets, bodies of water, parks and buildings (with textures). Users can optionally specify categories of places (i.e. restaurants, shopping, ...) they are interested in. Our system automatically determines the salience and importance of map elements using bottom-up vision-based image analysis and top-down web-based information extraction techniques. It then generates a map that emphasizes the most important elements, using a combination of multiperspective rendering to increase visibility of streets and landmarks, and cartographic generalization techniques such as simplification, deformation, and displacement to emphasize landmarks and de-emphasize less important buildings.

2 Tourist Map Design

In order to design more effective tourist maps we follow the approach of Agrawala and Stolte [August 2001] and begin by analyzing prior work on mental representations of cities [Lynch 1960; Sorrows and Hirtle 1999] as well as collections of the best hand-designed maps [Southworth and Southworth 1982; Tufte 1990; Holmes 1991] and cartographic textbooks [Kraak and Ormeling 1996]. From this analysis we extract the importance of the elements in a tourist map and identify a set of principles for rendering useful tourist maps.

2.1 Important information

In his classic book, *The Image of the City*, Kevin Lynch [1960] identifies five elements that people use to form mental representations of cities: *landmarks*, *paths*, *districts*, *nodes* and *edges*. Indeed, these elements have a number of properties that make them essential in navigational tasks and in the general understanding of a new environment.

Landmarks. Large physical objects often act as reference points in the environment. While our work focuses on buildings, other objects such as bridges and mountains can also serve as such landmarks. The principal characteristic of landmarks is that they are uniquely memorable in the context of the surrounding environment. Sorrows and Hirtle [1999] consider three subcategories of landmarks:

- *Cognitive landmarks* are semantically meaningful because of either their cultural or personal significance. For example, the apartment of a famous author is culturally significant, while a particular restaurant may be personally significant.
- *Visual landmarks* are buildings that are distinctive because of their appearance characteristics, such as their color or shape.
- *Structural landmarks* are memorable because of their spatial location. Examples include buildings located at decision points such as street intersections, bus stations or around town squares.

Building of all three subcategories should be included in well designed tourist maps. While cognitive landmarks characterize a place or reflect the personal interests of a viewer, visual and structural landmarks are essential for navigation. Such landmarks pro-

vide additional context for navigating through unfamiliar areas and enable users to localize themselves within the surrounding environment [Deakin 1996; Michon and Denis 2001].

Paths. Roads, walkways, transit lines, canals, railroads, etc. are paths through the environment on which people may freely assemble, interact, and move about. In this work we focus on roads, since they are the predominant paths for urban navigation.

Districts. Some areas of a city may have common identifying characteristics along a variety of dimensions including building type, building use, types of inhabitants etc. Neighborhoods such as Chinatown, Little Italy and the Mission are examples of such districts.

Nodes. In many cities, parks, town squares, beaches and busy intersections are points where people tend to congregate. Such nodes are particularly relevant for tourists because they are good places to mingle with the local population.

Edges. Many cities contain elements such as rivers, city walls, and roads that serve as linear breaks in continuity between regions. Such edges can be barriers which close one region off from another, or seams along which two regions are related and joined together.

Edges are often defined by the boundaries of the other four city elements. For example a highway serves as a barrier but also as an important path. Our system does not explicitly treat edges, but handles them implicitly by working with the other four elements.

2.2 Rendering

Multiperspective Maps

Cartographers often use multiple viewpoints and perspectives in order to increase the visibility of important elements of the map. A common approach is to combine an orthographic, top-down, plan view for the streets and ground plane, with either an oblique or perspective projection for the buildings. The top-down view of the ground plane eliminates foreshortening distortions and thereby allows users to better understand the layout of the road network. Cartographers choose oblique or perspective views for the landmarks to ensure that the street-side facades of the buildings are clearly visible.

Oblique and perspective projections have complementary advantages. With an oblique projection the size and shape of the building does not depend on the distance to the viewpoint. Agrawala et al. [June 2000] point out that setting the oblique projection image plane parallel to the ground plane ensures that the building footprints retain their true shape and thereby facilitate some size and area comparisons. However, the lack of perspective cues can also make the buildings appear unrealistic. Using a perspective projection for the buildings prevents this problem, but makes it difficult to visually compare size and area of buildings.

Generalization

Mapmakers use a variety of cartographic generalization techniques including simplification, displacement, deformation and selection to improve the clarity of the map and to emphasize the most important information, while preserving spatial relationships between map objects [Monmonier 1993; MacEachren 1995; DiBiase et al. 1992]. In this work, we consider the generalization techniques mapmakers apply to buildings in hand-drawn tourist maps and show how these techniques improve map usability.

Simplification. Mapmakers use simplification to de-emphasize less important buildings or to reduce the complexity of scaled down buildings and thus avoid rendering artifacts.

Displacement. Artists often widen roads in order to emphasize the paths in an unknown area. As a consequence, the space available for buildings within the block is limited. A common technique is to

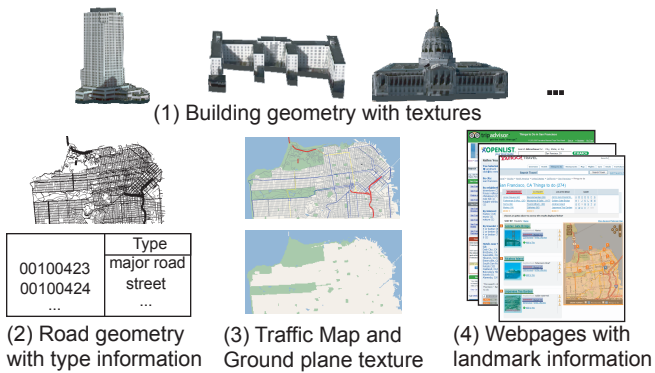


Figure 2: The input data to our system.

displace the buildings in order to avoid overlaps with neighboring streets or buildings.

Deformation. Mapmakers often increase the size of important landmarks while decreasing the size of less relevant buildings.

Removal. Handdrawn maps often remove less relevant buildings to reduce the complexity of maps and to gain space to emphasize more important buildings.

3 System Overview

We have developed an automated system for generating tourist maps based on the design principles presented in Section 2. As shown in Figure 2, the input data to our system consists of a complete geometric model of a city in lat/lon coordinates, including 3D building geometry with textures as well as roads geometry with each road being categorized as a street, major road, arterial, ramp or highway. A traffic map separates the streets into four discrete levels of traffic. A ground plane image segments the city into waterways, parks and ground. Finally, as explained in Section 4.1, our system parses a set of webpages¹ from which we extract semantic information about buildings including their category (museum, restaurant etc.), neighborhood, and a user ranking for each building.

From this data our system precomputes building footprints and block boundaries, where a block refers to the smallest area delimited by roads. We determine the building footprints by projecting the building triangles on the ground plane and tracing out the outline of the building in pixel space. We compute the block boundaries by triangulating the road network and applying a flood-fill over the triangles stopped by the street segments.

After preprocessing we identify landmarks, paths, districts, and nodes, as well as importances for each of these elements (Section 4). To generate a tourist map, a user specifies a city, or a region within a city, and the landmark categories of personal interest (i.e. restaurants, shopping, etc.). The system then designs and renders a tourist map that covers the specified region and highlights the personal points of interest (Section 5).

4 Computing Importance

Our system automatically extracts landmarks, paths, districts and nodes and their relative importance by constructing a set of features from the underlying city data. Following the approach of Sorrows and Hirtle [1999], we divide the features into three categories; *semantic*, *visual* and *structural*. The semantic features are computed via web-based information extraction techniques, while the visual and structural features are based on low-level analysis of the city geometry, textures, and ground plane image. Every feature is associated with a score, where a high score indicates a high importance.

¹www.openlist.com, www.tripadvisor.com, and travel.yahoo.com

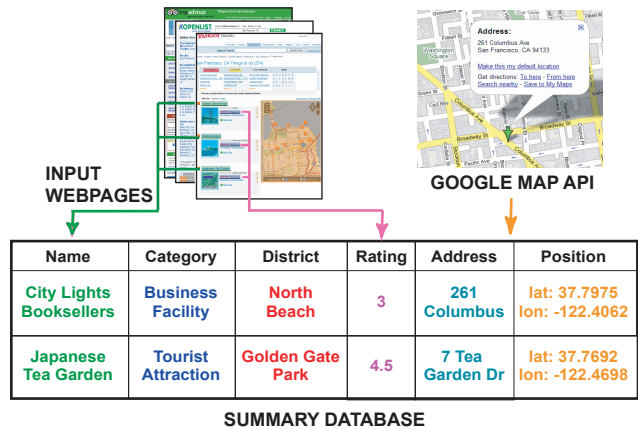


Figure 3: We extract landmark attributes from three input websites. When the websites only provide the address of a landmark we look up the exact lat/lon position by querying the Google Map API.

We compute the overall importance for each map element as a weighted sum of the feature scores. While users must pick the landmark categories they are most interested in, they can also manually set the weight for each feature in the weighted sum. Alternatively we provide a set of experimentally chosen default weights.

While Raubal and Winter [2002] and Klippel and Winter [2005] develop a similar feature-based approach for identifying important landmarks, they assume that the semantic features are given as part of the input. In contrast we extract semantic information from websites. We also consider how the importance of the five basic map elements (landmarks, paths, etc.) effect one another while they compute importance of each of the five elements independently.

4.1 Landmarks

We base our landmark features on the work of Appleyard [1969] who conducted a survey asking people what buildings they remembered along a road. From this survey he suggested a set of salient landmark features that we incorporate into our system.

Semantic Features

Identifying semantically meaningful features requires human knowledge of the cultural or personal significance of the landmark. While such semantic meaning cannot be determined from the geometric city data alone, the Internet provides a vast collection of semantically structured data from which we extract the appropriate features. A number of recent techniques exploit the web to extract such semantics in the context of photo manipulation [Snavey et al. 2006; Hays and Efros 2007]. Similarly Tezuka and Tanaka [2005] apply text-mining techniques to extract landmark information from unstructured web documents.

We extend this idea to parse landmark information from three travel-related websites. As shown in Figure 3 we extract the following attributes: name, lat/lon position, address, district, category (tourist attraction, museum, business facility, shopping, restaurants etc.) and user ranking (0 to 5 stars). When the travel website does not contain the landmark position our system queries the Google Map interface with the building address to retrieve the position. We use the lat/lon position to associate each landmark with the corresponding building geometry from our city data set. Finally, we take the average user ranking of the landmark across the three websites as its semantic score. Figure 6(left) shows the most important semantic landmarks in an area.

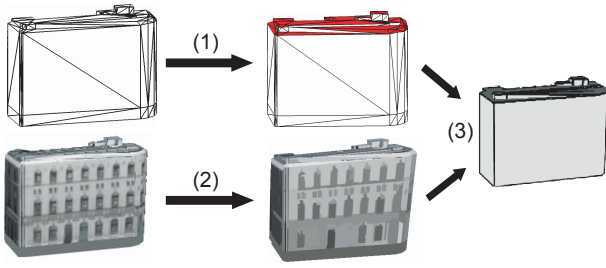


Figure 4: Extraction of building color. (1) We first distinguish between roof and wall triangles, (2) segment the texture and average the color of every segment, (3) and then we set the wall color to the color of the brightest segment in the wall area.

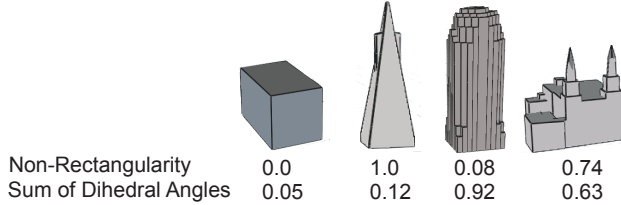


Figure 5: Shape complexity features. The first rectangular building has low complexity by both measures. The second building is unusual because it is highly non-rectangular, but it also has a low sum of dihedral angles. The third building is fairly rectangular but also has many corners and therefore has high shape complexity. The last building is high shape complexity by both measures.

Visual Features

Buildings with distinctive visual features (color, shape, etc.) that differ significantly from the appearance of other buildings in the local area are likely to be remembered as landmarks. We distinguish three types of visual characteristics that strongly influence the saliency of a building: *facade color*, *shape complexity* and *building height*. We first define a set of features that quantify the presence of these characteristics in a building and then explain how we associate scores with these features.

Facade color. Some buildings are memorable because their facades are colored differently from the surrounding buildings. As shown in Figure 4, to extract the wall color of a building, we first separate building geometry into roof and wall triangles. A building triangle i with normal n_i is labeled as a wall, if $n_i \cdot (0, 0, 1) < 0.5$, and as a roof otherwise. Next, for all wall triangles we apply Felzenszwalb’s [2004] color-based segmentation algorithm on the corresponding texture map. We approximate the color of each segment by its average color to smooth out small color irregularities and noise in the texture. Finally, to avoid dark shadows in the texture, we set the facade color to the the brightest color in the segmented texture, where we consider only the segments that have an area larger than 5% of the total wall area. The color value is the L component of the facade color in the LAB color space.

Shape complexity. Many buildings have a simple, regular, rectangular shape. The more a building deviates from this standard shape, the more salient it becomes as a landmark. We measure the shape complexity using two features. The first feature measures the rectangularity of a building as the ratio of the volume of the building’s bounding box to the true volume of the building [Mirtich 2005]. The second feature measures the angle variation of the building triangles and is computed as the sum of absolute dihedral angles weighted by the length of the triangle edge. Figure 5 illustrates the difference between these two shape complexity features.

Building height. Taller buildings are more likely to be visible from further away. Therefore we use building height as a visual feature

with taller buildings getting higher value.

In the context of visual features, a building only becomes a landmark if it differs significantly from its neighboring buildings. To compute the score of each visual feature, we thus estimate the deviation from the median using a measure proposed by Nothegger et al. [2004]:

$$score = \frac{|x - med(x)|}{med(|x - med(x)|)},$$

where x is the individual measure and $med(x)$ denotes the median. To compute the median we define the local neighborhood of each building to include all other buildings within 50m. Figure 6(middle) shows the most important visual landmarks in an area.

Structural Features

Several cognitive psychology studies have shown that travelers are more likely to take note of buildings located at road intersections and around town squares as they learn to navigate a new city [Denis et al. 1999; Lovelace et al. 1999; Michon and Denis 2001]. The spatial locations of these buildings makes them especially relevant for navigation and therefore we identify both of these types of structural landmarks.

Buildings at intersections. Buildings at important intersections are more prominent than others. We compute a score for each intersection as the sum of the importance scores of all roads (see Section 4.2 on computing road importance) meeting at that intersection. Thus, we favor intersections of important roads or places where many roads meet. Finally, we transfer the importance of an intersection to the nearby buildings using a Gaussian weighting factor based on the distance between the intersection and the building. For each building, we accumulate influences from all nearby intersections.

Buildings around squares. To identify buildings located around town squares we first extract the squares and their importance as explained in Section 4.4. The importance of the square is then transferred to the buildings nearby analogously to the intersection importance.

Figure 6(right) shows the most important structural landmarks in an area.

4.2 Roads

We use semantic and structural features to identify the important roads. While semantic features are the same for all maps, the structural features depend on the points of interest chosen by the user.

Semantic Features

Previous work on the selection of important roads has focused on inferring semantic importance indirectly from the topology and geometry of the roads [Mackaness and Beard 1993; Thomson and Richardson 1995; Thomson and Richardson 1999; Jiang and Claramunt 2004]. In contrast, our system directly accesses this semantic information from the input data or the web.

Categories. Roads are often classified into a discrete set of categories according to their usage. For example, each road in our input data is categorized as either a street, major road, arterial, ramp or highway. The respective scores for each categories are 0.25, 0.5, 0.75, 1.0, 1.0.

Traffic. Customary travel is one of the strongest influences in the importance of a road [Lynch 1960]. Major access lines such as the Bay Bridge or the Presidio Parkway in San Francisco are key features of the mental map. We retrieve the information about road use from a traffic map (travel.yahoo.com) that classifies road segments into four discrete levels of traffic according to their usage during rush hours. The traffic score of a road segment corresponds to its normalized traffic level.

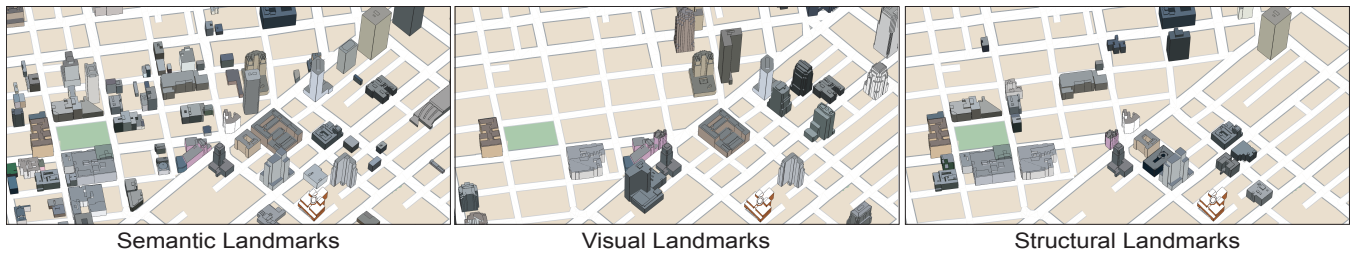


Figure 6: Comparison of semantic, visual and structural landmarks in an area.

Structural Features

We specifically design the structural features to select for roads that facilitate navigation to and around interesting areas for tourists. Our roads are represented as polylines and we compute importance scores for each linear segment of the road. The segment importance indicates how relevant a road is locally and influences the importance of nearby landmarks, as explained in Section 4.1.

Landmark proximity. For tourists, the roads closest to their personal points of interest are especially important because they facilitate navigation. To compute this feature, we first estimate the semantic importance for the landmarks chosen by the user as points of interest (restaurants, shopping, etc.) as described in Section 4.1. We then subsample every road segment at 2m intervals and transfer the semantic importance of nearby landmarks to these road sample locations using a Gaussian weighting factor based on the distance between the road sample and the landmark location. At each road sample point we accumulate the influence of all nearby landmarks. Finally, for each linear segment of a road we compute the average score of its subsampled points.

Roads connecting landmarks. Since tourists often want to circulate between different sights of the city, it is helpful to emphasize the connecting roads. We query the Virtual Earth map control API (<http://dev.live.com/virtualearth/sdk>) to obtain the route between each pair of landmarks. The score for each road segment is then computed as the number of times that segment is part of a landmark to landmark route. Unlike landmark proximity, this feature can increase the importance of roads that do not have many landmarks on them, but are often used to travel between interesting parts of the city.

4.3 Districts

Semantic Feature

In identifying the semantic features for each landmark we also extract the district or neighborhood the landmark belongs to (see Section 4.1). Districts usually do not have a hard or precise boundary and therefore we simply identify the district center, by searching for the location with the highest density of landmarks from that neighborhood. We determine this location by iteratively computing the mean location of all landmark positions in that district and rejecting outliers at each iteration if necessary.

4.4 Nodes

Semantic Feature

Some nodes such as squares, parks and lakes are also considered landmarks and we extract them by parsing travel websites in exactly the same way we identify semantic landmarks (see Section 4.1). However this extraction process provides only a point location for the node. To compute the 2D extent of squares we expand the point location to the entire surrounding block. Similarly we use a color-based flood-fill on the ground plane texture seeded at the point locations of parks and lakes to determine their 2D extents. The score of a node, as for the semantic landmark feature, corresponds to the average user rating indicated in the input websites.

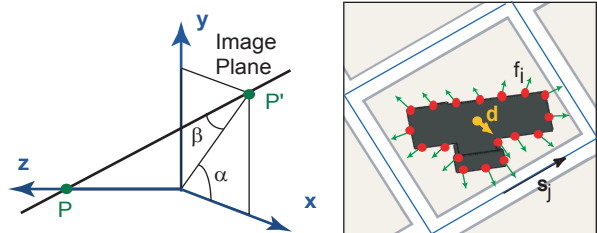


Figure 7: (left) Oblique projection parameters. The projection angle β is the angle between the projector line and the image plane. The orientation angle α is the angle in the image plane between the x-axis and the projector. (right) Computing the orientation vector \mathbf{d} . For each footprint point f_i we compute a weighted sum (green vectors) over all the block segment normals where the closest segments are given higher weight. We then sum these green vectors to obtain \mathbf{d} .

5 Rendering

5.1 Multiperspective Maps

Multiperspective maps have the advantage of depicting roads and the visually salient facades of buildings in a single view. To create such multiperspective maps we render the ground plane using an orthographic projection and either an oblique or perspective projection for the landmarks.

Oblique Projection

As shown Figure 7, the oblique projection is defined by two parameters β and α that specify the direction of the projector lines. In our application the image plane is parallel to the ground plane and β controls the amount of foreshortening of the building facades, while α controls which facades of the buildings are visible. Because the image plane is orthogonal to the ground plane the buildings remain correctly orientated with respect to the streets – the oblique projection does not change the orientation of the building footprint.

We build a different oblique projection for each building such that the building facades closest to the nearby roads are visible. When more than two facades face a street, we favor the facades facing the most important streets. More specifically, we compute an orientation vector \mathbf{d} for each building that represents the direction to the closest roads. As shown in Figure 7 we subsample the building footprint to form the points f_i for $i = 1..M$ and break the block boundary into a set of segments s_j for $j = 1..N$ with normals \mathbf{n}_j . Then

$$\mathbf{d} = \frac{1}{MN} \sum_i \sum_j W(f_i, s_j) \mathbf{n}_j$$

is a weighted sum of the segment normals. The weight is given by

$$W(f_i, s_j) = k_j \left(1 - \frac{D(f_i, s_j)}{\max D(f, s)} \right),$$

where k_j is the importance score of the road segment s_j (see Section 4.2 for road importance), $D(f_i, s_j)$ is the distance between the footprint point f_i and the block boundary segment s_j , and $\max D(f, s)$ is computed over all footprint points and boundary seg-

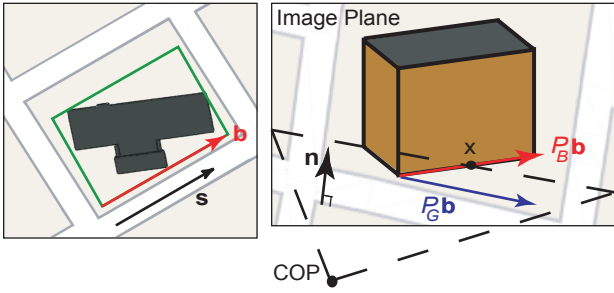


Figure 8: (left) The building footprint is best aligned to street segment s . We compute a bounding box for the footprint in the direction of s and then identify the bounding box edge \mathbf{b} closest to s . (right) To rotate the buildings we first construct the plane in world space that passes through the center of projection (COP) of P_B and the point x , and intersects the image plane parallel to $P_G \mathbf{b}$

ments. Thus, boundary segments that are closest to the footprint points are given the highest weights.

Finally we use \mathbf{d} to compute the projection parameters as

$$\alpha = \arctan \frac{d_y}{d_x} \quad \text{and} \quad \beta = \frac{\pi}{2} - c \|\mathbf{d}\|.$$

We set α to the orientation of \mathbf{d} so that the street-side building facades are visible. We set β proportional to the length of \mathbf{d} and use the scale factor c to control the amount of foreshortening.

Perspective Projection

Rendering the buildings using a perspective projection provides better depth cues and less distortion than using an oblique projection. However, the perspective projection also creates misalignments in translation, rotation and scale between the buildings and the roads that are rendered using the top-down orthographic projection of the ground plane.

To correct these misalignments we first identify the road segment that each building is most aligned with. We initially work in world space and project each vertex of the building footprint geometry onto each road segment adjacent to the building. We then choose the road segment s with the longest range of projected points (see Figure 8(left)). Next we compute a vector \mathbf{b} that describes the orientation of the building with respect to s . We fit a bounding box oriented in the same direction as s to the building footprint. Then, \mathbf{b} is the edge of the bounding box closest to s .

Suppose that P_B and P_G designate the building projection matrix and the ground projection matrix respectively and that x is the midpoint of \mathbf{b} . To correct for translational misalignment, we compute a translation matrix T in building space such that $P_B T x = P_G x$, and apply T to all vertices of the building geometry.

To correct for rotational misalignments we compute a rotation matrix R_z about the z -axis of building space such that $P_B R_z \mathbf{b}$ and $P_G \mathbf{b}$ are aligned. Note that we cannot set $R_z = P_B^{-1} P_G$ because in the general case $P_B^{-1} P_G$ will not be a rotation about the z -axis. As shown in Figure 8, our approach is to construct a plane in world space that passes through the center of projection (COP) of P_B and the point x , and intersects the image plane parallel to $P_G \mathbf{b}$. To build the plane we compute a world space vector \mathbf{b}' such that $P_B \mathbf{b}' = P_G \mathbf{b}$, by taking two points on $P_G \mathbf{b}$, giving them the same z coordinate and then transforming them through P_B^{-1} to put them in world space. Then the normal to the plane is given by $\mathbf{n} = (x - \text{COP}) \times \mathbf{b}'$. Since this plane passes through the COP any edge that lies in it must be parallel $P_G \mathbf{b}$. Thus, we construct the rotation matrix R_z that rotates \mathbf{b} about the point x such that \mathbf{b} lies in the plane. That is, $R_z \mathbf{b} \cdot \mathbf{n} = 0$. Solving for the rotation angle θ we obtain

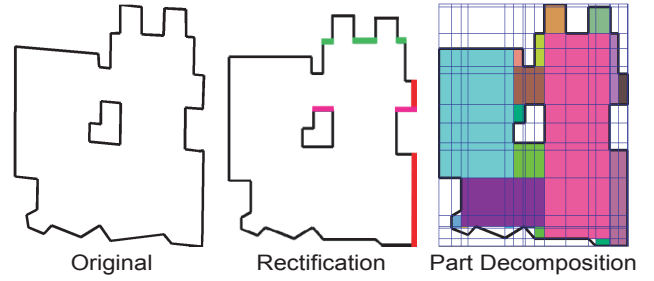


Figure 9: Building Simplification. (1) The original footprint. (2) The building footprint after rectification. Colored edges indicate the groups of edges that are snapped to the same horizontal or vertical cell boundary. (3) The building footprint after fitting a grid to it. Using a 2D flood fill we fit rectangles to the full cells.

$$\theta = \arcsin \frac{-b_z n_z}{\sqrt{c^2 + d^2}} - \arctan \frac{d}{c},$$

$$\text{where } c = b_y n_x - b_x n_y \quad \text{and} \quad d = b_x n_x + b_y n_y.$$

After rotating the buildings to re-align them with the roads we scale down buildings if their footprint is larger than their block area. We apply P_B to the building footprint vertices and P_G to the block boundary vertices. Once all vertices are in image space, we perform a binary search to find the scale factor that will reduce the area of the building footprint to fit within the block area. We scale the buildings about x , because x is at an appropriate distance from the block boundaries and we would like to keep this distance fixed.

5.2 Generalization

Building simplification

Mapmakers often use simplification as a tool to direct a viewer's attention to buildings with more geometric detail by reducing the visual complexity of less important buildings. While automated building simplification is a well-studied area of cartography, most approaches focus on simplifying 2D building footprints [Staufenbiel 1973; Sester 2000]. More recent techniques for simplifying buildings in 3D are designed to find and eliminate small volumetric features such as protrusions on the surface model [Kada 2002; Thiemann 2002; Forberg 2007].

Our simplification algorithm is based on the method of Forberg [2007] who focuses on the the problem of simplifying buildings composed of axis-aligned orthogonal planes. She searches for planar building facets that are parallel and located near one another and then shifts these facets towards each other until they lie on a single plane. We extend this work to handle some non-orthogonal buildings. Our approach proceeds in 3 phases; *rectification*, *part decomposition*, as shown in Figure 9, and *facet shifting*, as shown in Figure 10. We assume that the buildings are vertical extrusions of their footprint and therefore most of our processing is done on the 2D building footprints.

Rectification. The rectification phase is designed to orthogonalize buildings by reorienting their walls to the principal orientations of the building footprint. To compute the principal axes we bin the building footprint edge segments by their orientation. We then identify the bin with the largest number of segments weighted by their length and fit a box to the building with the average orientation of the segments in the bin. All subsequent phases treat these principal axes as the x - and y -axes of the building. We further rectify building footprint edges that are almost parallel (i.e. within 10°) to the principal axes by snapping them to these axes. Similarly we group together parallel footprint edges that are near one another and clamp them to lie on the same line in order to better preserve the structure of the building during simplification.

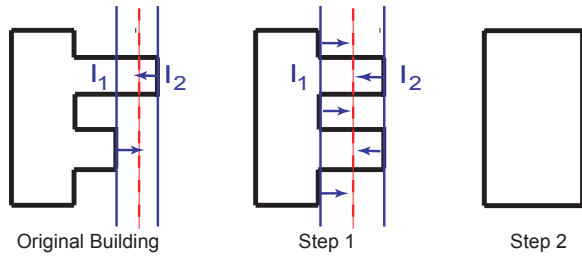


Figure 10: Facet Shifting. In each step the closest parallel facets of the building, l_1 and l_2 , are shifted to lie on the same line

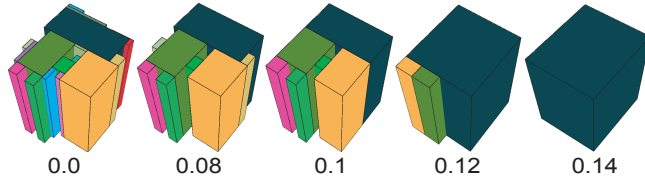


Figure 11: Five steps of building simplification. The coefficients below each step indicate the relative amount of simplification.

Part decomposition. We decompose the building into rectangular sub-parts by first embedding it in a non-uniform grid. We generate grid cell boundaries along the principal x -, y - and z -axes at every footprint and roof vertex in the model. We then classify the cells as being fully inside, partially inside, or outside the building. To classify a cell we uniformly subsample it and check whether each sample lies inside the building using a point in polyhedron test. If more than 5% of the samples are inside the building it is marked as partially inside and if more than 95% are inside it is marked as fully inside. Finally, we group together grid cells marked as fully inside into larger rectangular subparts using a greedy flood-fill algorithm.

Facet shifting. To simplify buildings we iteratively apply Forberg's [2007] facet shifting algorithm. We first identify the two consecutive grid lines in our part decomposition, l_1 and l_2 , with the smallest spacing in either the x or y direction. As shown in Figure 10, we then replace these two lines by a new line at the average x or y location and shift the corresponding facets to this new line. In most cases this process causes the boxes delimited by l_1 and l_2 to disappear and increases the size of neighboring boxes.

At each iteration of this algorithm we quantify the amount of simplification as the volume of the model eliminated by facet shifting normalized by the volume of the initial, unsimplified orthogonal model. Figure 11 shows the result of iteratively simplifying an example building and Figure 13 shows simplification applied to a larger area.

This algorithm cannot simplify buildings with significantly inclined roofs or walls such as San Francisco's Transamerica pyramid. However, such non-orthogonal buildings are uncommon in most cities. In our San Francisco data set, 7% of the 6227 buildings are non-orthogonal and remain unsimplified. Moreover, such buildings tend to be visually distinctive and therefore should be treated as important visual landmarks that are rendered without simplification.

We identify non-orthogonal buildings by computing the volume of the partially inside cells over the total building volume. If this measure is below 10%, we drop the partially inside cells and treat the building as orthogonal. Otherwise we mark the building as irregular and do not simplify it. Thus, our approach orthogonalizes buildings with small deviations from orthogonality, but does not simplify large irregular buildings.

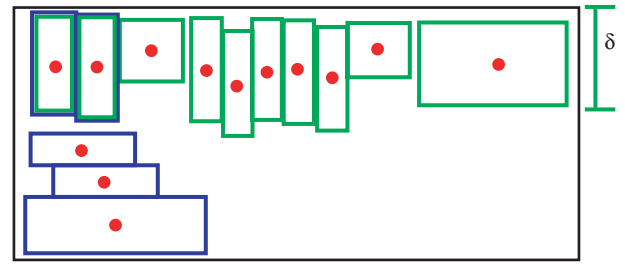


Figure 12: Alignment groups. Buildings that are within a distance δ of an edge of a block boundary form an alignment group. The green alignment group is within δ of the top block edge and the blue group is within δ of the left block edge. Buildings like the ones in the top left can belong to more than one alignment group.

Building Layout

Mapmakers often widen roads in order to emphasize the paths in an unknown place and prevent buildings from occluding the roads in dense areas. As a consequence, the space available for buildings in the block is limited. A common approach is to use map generalization to adjust the buildings to fit within the limited space [McMaster and Shea 1992; Ware and Jones 1998; Weibel and Jones 1998; Bader 2001; van Kreveld 2001].

We follow the approach of Ware and Jones [1998] and generalize the buildings within a block using simulated annealing. While their work only treats the displacement of buildings, we include scaling and removal of buildings to allow for greater flexibility in generalization. Like any simulated annealing-based technique we must define three procedures in order to run the optimization: *initialization*, *perturb* and *scoring*.

Initialization. Widening the roads reduces the size and changes the boundary of each block. Thus, to initialize the optimization, we first determine the new block boundaries by re-triangulating the widened road network and applying a flood-fill over the triangles stopped by the streets. For each building, we determine a new approximate position within the reduced block by computing the orientation vector to the center of the original block and scaling it by the best uniform scale mapping between the original block and the reduced size block. Finally, we compute alignment groups for a block that contain all the buildings aligned to the same street as shown in Figure 12.

Perturb. The perturb function for the search randomly picks a building to alter and randomly chooses a generalization operator for that building. We have implemented three operations: displacement of the building parallel or perpendicular to the direction of the alignment group, scaling of the building, and removal and reinsertion of the building.

Scoring. The layout scoring function evaluates each building on a number of criteria. More specifically, we penalize a building if (1) it falls outside the block boundary, (2) it moves far away from any street it is aligned with, (3) it falls out of alignment with other buildings in the same alignment groups, (4) its position ordering within its alignment groups changes, (5) it is reduced in scale significantly and it has high importance, (6) it is scaled differently from its neighbors, (7) it overlaps other buildings, (8) it is removed from the map.

If conflicts in generalization, such as building overlaps, cannot be resolved by the annealing process, the resulting map will contain artifacts. In general the system chooses a suitable compromise, yet our automated generalization methods cannot compete with the best works of highly proficient map artists. Nevertheless, we believe that our visualizations surpass in quality many existing tourist maps in use today.

Figure 13 shows block generalization applied to an area around



Figure 13: Generalization. (Top) Original block. (Middle) Importance scores for the buildings, where the red saturation of a building is proportional to its importance. (Bottom) Generalized blocks.

Union Square in San Francisco. The simplification coefficients for the buildings are inversely proportional to their importance score (see Section 4.1).

5.3 Non-Photorealistic Rendering

Mapmakers often emphasize landmarks by marking their contour lines and attenuating the building color. Contour lines convey information about the shape and help a viewer to better identify the building’s geometry. We determine the set of contour lines by identifying the edges between any two neighboring triangles that have a dihedral angle of more than $\pi/4$.

Artifacts in the building texture such as shadows can be very distracting for a viewer and make the building harder to recognize. To counteract this effect, our system can render the walls and the roofs using a uniform color. After extracting the wall and roof colors, as explained in Section 4.1 we further saturate the wall colors to make them more distinctive and we brighten the roof colors to make the buildings appear as if they are lit from above. To de-emphasize non-landmark buildings we render them in light gray.

5.4 Labeling

For tourist maps to be usable and to allow navigation, we must label important map elements including landmarks, roads, nodes (lakes, parks, squares) and districts. Automated placement of labels and line features on maps is a well-studied problem in automated map design [Imhof 1975; Edmondson et al. 1996; Agrawala and Stolte

August 2001; Vollick et al. 2007]. A common approach is to search the space of possible labelings of the map to find an optimal layout. We build on the simulated annealing approach of Agrawala and Stolte [August 2001].

We label the map elements in four stages starting with nodes, and then the landmarks, roads and districts in consecutive order. Breaking the search into stages in this manner, reduces the dimensionality of our search space and significantly improves convergence. Our system extends the previous labeling techniques by including the importance values of the map elements in the label layout scoring function. This approach ensures that the most important elements will be labeled in the optimal way, while the labels of less important elements might be placed further away or even removed.

6 Results

Examples of several tourist maps generated using our system are presented in Figures 1, 16 and 17. The input data consists of a set of 6227 buildings of downtown San Francisco. We identify 1257 of these as semantic landmarks by parsing three input websites (see Section 4.1).

Figure 15 shows the advantage of widening the roads and generalizing the buildings for the area around Grace Cathedral in San Francisco. In the left image California and Clay St are not visible. After widening the roads and generalizing the buildings, the two occluded streets become visible while the overall structure of the block is preserved.

The maps of Figure 16 show a set of distinctive landmarks in San Francisco for different categories of interest. The lower maps correspond to the categories: tourist attraction (left), restaurants (middle) and shopping (right) and the upper image is a combination of the most important landmarks in San Francisco disregarding their category. Note that Figure 1 shows a labeled version of the restaurant map. In all of these maps we emphasize points of interest by drawing them in color with dark labels, while other salient buildings and their labels are drawn in light gray. In order to increase the visibility of the landmarks and the ground plane, we render the buildings using a perspective projection and the roads with an orthographic projection. This map only displays the most important roads. The width of a road and the color of its label depend on the road category.

Figure 17 shows a close-up view of the buildings along Market Street. We use an oblique projection for the buildings which emphasizes the street-side facades of the buildings. In this case the facades facing Market Street are always visible since this is the street with the highest importance in that area. We label semantic landmarks and de-emphasize other buildings by making them semi-transparent.

Comparison with existing maps. In Figures 1, 14 and 16 we compare existing digital maps and hand-drawn maps of San Francisco to the tourist maps generated with our approach. Unlike the digital maps which show building footprints or simplified 3D building geometry directly texture mapped with satellite imagery, we select the relevant information for tourists and render more recognizable abstracted 3D views of the landmarks. In addition, we use cartographic generalization and multiperspective rendering to reduce the complexity of the map, emphasize the most important landmarks and facilitate navigation.

While our maps do not yet match the rendering quality of the best hand-designed maps created by expert mapmakers (see Figure 14(left)), our maps have the advantage of being able to adapt to changes in the city and to the interests of the user. A hand-drawn map must show all the landmarks that might be of interest to any viewer and it is therefore usually cluttered with buildings that occlude one another. The clutter also makes the labels difficult to

read. In contrast, our maps emphasize only the landmarks that a individual user is most interested in and is therefore more legible.

Timing information. Our tourist map generation system is comprised of several preprocessing steps followed by an authoring interface that generates the final maps based on the user's interests. In the preprocessing steps we compute the semantic, visual and structural features for the buildings, roads, districts and nodes. We also precompute the part rectification and part decomposition of the buildings as well as road independent weights for the oblique projection. Because these computations are performed once for all buildings in our data set, we have not focused on optimizing them and many of them are implemented in Matlab. While most of the preprocessing steps take on the order of minutes, a few take longer. The slowest requires computing the route between all pairs of semantic landmarks (a structural road feature) and takes about 24 hours. We rely on the Virtual Earth web-based service to query these routes.

Our map authoring interface loads this precomputed data and allows users to modify importance weights, set the viewing parameters and the landmark categories of interest. To display a map we generalize the buildings, perform multi-perspective rendering and then label the map. Generalizing a single block of buildings can take between a few seconds when the block contains very few landmark buildings, up to a few minutes when the block contains many landmark buildings. Multi-perspective rendering occurs in real-time. Labeling is relatively slow and proportional to the number of labels that must be placed in the map. While the map in Figure 16 contains hundreds of labels and takes about 40 minutes to label, the map in Figure 15 takes only a few seconds. However we believe that using a specialized data structures to speed up the spatial overlap computations as well as GPU accelerations can significantly reduce labeling time by 1 or 2 orders of magnitude [Hartmann et al. 2005; Ali et al. 2005].

7 Conclusions and Future Work

We have introduced a comprehensive system for automatically generating tourist maps. Our system exemplifies how a combination of web-based data mining, perception-guided geometry and image analysis, and advanced visualization techniques yields an effective tool for the automatic creation of tourist maps suitable for everyday use.

One of the greatest appeals of our approach is customization. Users can easily build their own maps according to their preferences and intentions for a particular visit to a city. In contrast to hand-drawn maps, our system makes use of up-to-date information and could thus even accommodate dynamic information such as road closures or specific opening hours of museums or restaurants.

While in this paper we focus on printed maps, one direction for future work is to produce interactive maps for mobile devices. For example, Zipf [2002] suggests interactively personalizing maps based on user preferences, cultural differences and GPS-based location. Such interactivity and personalization could further improve the effectiveness of automatically generated tourist maps.

8 Acknowledgments

We thank Eyal Ofek for suggesting the directions we have pursued in this work. We would also like to thank Eyal Ofek, Bill Chen and the Microsoft Virtual Earth team for providing access to the geometric city data. The closed form solution for the multiperspective rotation grew out of conversations with James O'Brien. This work is partially supported by a gift from Microsoft Virtual Earth, an Alfred P. Sloan Foundation fellowship and a National Science Foundation grant CCF-0643552.

References

- AGRAWALA, M., AND STOLTE, C. August 2001. Rendering effective route maps: Improving usability through generalization. *In Proc. SIGGRAPH*, 241–249.
- AGRAWALA, M., ZORIN, D., AND MUNZNER, T. June 2000. Artistic multiprojection rendering. *Eurographics Rendering Workshop*, 125–136.
- ALI, K., HARTMANN, K., AND STROTHOTTE, T. 2005. Label Layout for Interactive 3D Illustrations. *Journal of WSCG* 13, 1, 1–8.
- APPLEYARD, D. 1969. Why buildings are known: A predictive tool for architects and planners. *Environment and Behavior* 1, 2, 131.
- BADER, M. 2001. *Energy Minimization Methods for Feature Displacement in Map Generalization*. PhD thesis, Univ. of Zurich.
- DEAKIN, A. 1996. Landmarks as navigational aids on street maps. *In Proc. CAGIS* 23, 1, 21–36.
- DENIS, M., PAZZAGLIA, F., CORNOLDI, C., AND BERTOLO, L. 1999. Spatial discourse and navigation: an analysis of route directions in the city of venice. *Applied Cognitive Psychology* 13, 2, 145–174.
- DIBIASE, D., MACÉACHREN, A., KRYGIER, J., AND REEVES, C. 1992. Animation and the role of map design in scientific visualization. *In Proc. CAGIS* 19, 4, 201–214.
- EDMONDSON, S., CHRISTENSEN, J., MARKS, J., AND SHIEBER, S. 1996. A General Cartographic Labelling Algorithm. *Cartographica: The Int. Journal for Geographic Information and Geovisualization* 33, 4, 13–24.
- FELZENSZALB, P., AND HUTTENLOCHER, D. 2004. Efficient graph-based image segmentation. *IJCV* 59, 2, 167–181.
- FORBERG, A. 2007. Generalization of 3d building data based on a scale-space approach. *In Proc. ISPRS* 62, 2, 104–111.
- HARTMANN, K., GÖTZELMANN, T., ALI, K., AND STROTHOTTE, T. 2005. Metrics for Functional and Aesthetic Label Layouts. *Int. Symp. on Smart Graphics* 5, 115–126.
- HAYS, J., AND EFROS, A. 2007. Scene completion using millions of photographs. *ACM Trans. Graph. (Proc. SIGGRAPH)* 26, 3, 4.
- HOLMES, N. 1991. *Pictorial Maps*. Watson-Guptill.
- IMHOF, E. 1975. Positioning names on maps. *The American Cartographer* 2, 2, 128–144.
- JIANG, B., AND CLARAMUNT, C. 2004. A structural approach to the model generalization of an urban street network. *GeoInformatica* 8, 2, 157–171.
- KADA, M. 2002. Automatic generalisation of 3D building models. *In Geospatial Theory, Processing and Applications* 34, 4.
- KLIPPEL, A., AND WINTER, S. 2005. Structural Saliency of Landmarks for Route Directions. *In Proc. COSIT 2005*, 347–362.
- KRAAK, M., AND ORMELING, F. 1996. Cartography. Visualization of spatial data. *Harlow: Longman*.
- LOVELACE, K., HEGARTY, M., AND MONTELLO, D. 1999. Elements of good route directions in familiar and unfamiliar environments. *In Proc. COSIT '99*, 65–82.
- LYNCH, K. 1960. *The Image of the City*. MIT Press.
- MACÉACHREN, A. 1995. *How Maps Work: Representation, Visualization, and Design*. Guilford Pubn.
- MACKANESS, W., AND BEARD, K. 1993. Use of graph theory to support map generalization. *In Proc. CAGIS* 20, 4, 210–221.
- MCMMASTER, R., AND SHEA, K. 1992. *Generalization in digital*

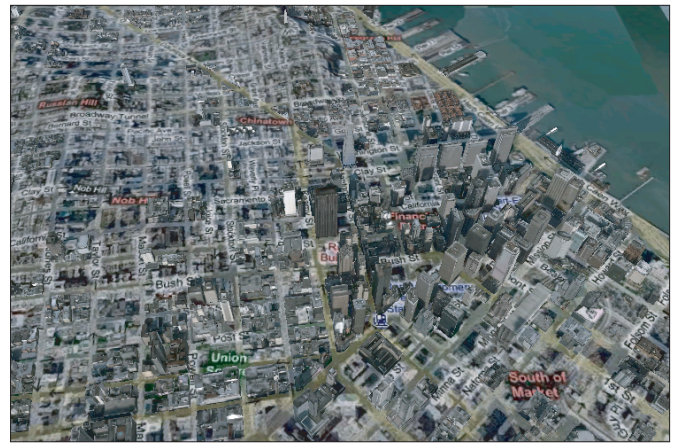


Figure 14: Comparison of our results with an existing manual and automated approaches. (left) Compared to a hand-drawn maps, our map can adapt to the interest of a user and changes in the city. Note that this map is from a different viewpoint than our result map shown in Figure 16. (right) Unlike the digital maps, we select the relevant information for tourists and render more recognizable views of the landmarks. In addition, we use cartographic generalization techniques to reduce the complexity of the map and multiple perspectives to increase the visibility of landmarks and roads.



Figure 15: Generalization applied on the area around the Grace Cathedral in San Francisco. After widening the streets and generalizing the buildings the occluded California and Clay St in the right image become visible (left).

cartography. Association of American Geographers.

MICHON, P., AND DENIS, M. 2001. When and why are visual landmarks used in giving directions? *In Proc. COSIT '01* 2205, 292–305.

MIRTICH, B. 2005. Fast and Accurate Computation of Polyhedral Mass Properties. *Graphics Tools: The JGT Editors' Choice*.

MONMONIER, M. 1993. *Mapping It Out: Expository Cartography for the Humanities and Social Sciences*. Univ. of Chicago Press.

NOTHEGGER, C., WINTER, S., AND RAUBAL, M. 2004. Computation of the salience of features. *Spatial Cognition and Computation* 4, 2, 113–136.

RAUBAL, M., AND WINTER, S. 2002. Enriching Wayfinding Instructions with Local Landmarks. *GIScience: Second International Conference* 2478, 243–259.

SESTER, M. 2000. Generalization based on least squares adjustment. *In Proc. ISPRS* 33, 931–938.

SNAVELY, N., SEITZ, S., AND SZELISKI, R. 2006. Photo tourism: exploring photo collections in 3d. *ACM Trans. Graph. (Proc. SIGGRAPH)* 25, 3, 835–846.

SORROWS, M., AND HIRTLE, S. 1999. The nature of landmarks for real and electronic spaces. *In Proc. COSIT '99*, 37–50.

SOUTHWORTH, M., AND SOUTHWORTH, S. 1982. *Maps: A Visual Survey and Design Guide*. Little Brown and Company.

STAUFENBIEL, W. 1973. *Zur Automation der Generalisierung topographischer Karten mit besonderer Berücksichtigung grossmasstäbiger Gebäudedarstellungen*. PhD thesis, Univ. of

Hanover.

TEZUKA, T., AND TANAKA, K. 2005. Landmark extraction: a web mining approach. *spatial information theory. Lecture Notes in Computer Science* 3693, 379–396.

THIEMANN, F. 2002. Generalization of 3D building data. *In Geospatial Theory, Processing and Applications* 34, Part 4.

THOMSON, R., AND RICHARDSON, D. 1995. A graph theory approach to road network generalisation. *In Proc. ICC* 17, 3–9.

THOMSON, R., AND RICHARDSON, D. 1999. The 'good continuation' principle of perceptual organization applied to the generalization of road networks. *In Proc. ICC* 19, 1215–1223.

TUFTE, E. 1990. *Envisioning Information*. Graphics Press.

VAN KREVELD, M. 2001. Smooth generalization for continuous zooming. *In Proc. ICC* 20, 2180–2185.

VOLLIK, I., VOGEL, D., AGRAWALA, M., AND HERTZMANN, A. 2007. Specifying label layout style by example. *In Proc. UIST*, 221–230.

WARE, J., AND JONES, C. 1998. Conflict reduction in map generalization using iterative improvement. *GeoInformatica* 2, 4, 383–407.

WEIBEL, R., AND JONES, C. 1998. Computational Perspectives on Map Generalization. *GeoInformatica* 2, 4, 307–314.

ZIPF, A. 2002. User-Adaptive Maps for Location-Based Services (LBS) for Tourism. *Information and Communication Technologies in Tourism*, 329–338.



Figure 16: Semantic landmarks in San Francisco. (top) A general-purpose map showing the most important landmarks across all semantic categories. (bottom) Tourists with differing personal interests can adapt the map to show only (left) tourist attractions, (middle) restaurants, or (right) shopping.



Figure 17: Market Street. Buildings are rendered with an adaptive oblique projection that ensures that street-side facades are visible so that tourists can easily recognize them as they walk down the street.