

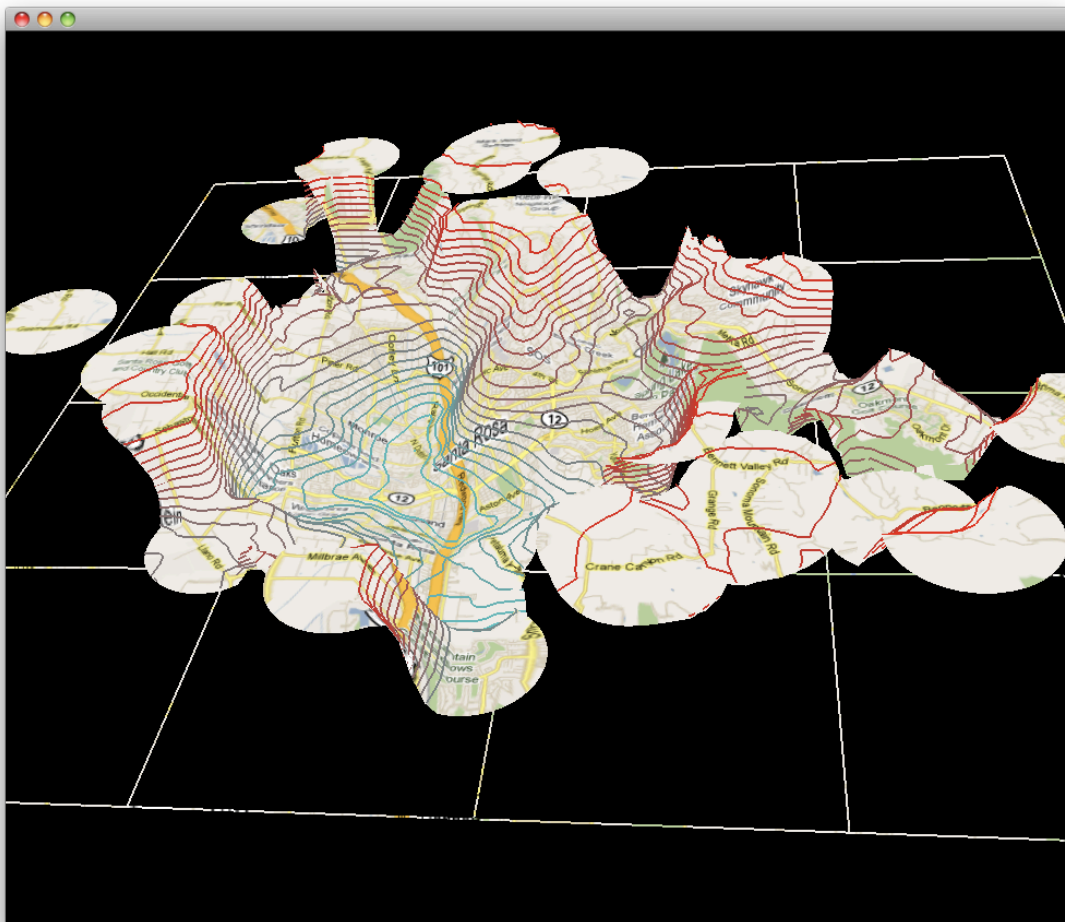
# Information Terrains

A Technique for Visualizing Information that has a Geographic Correlation

Eric Biggs    CS 284    Final Project    Fall 2009

## Introduction

I have implemented a technique to help visualize geographic information by using the third dimension over a 2 dimensional map to convey the information. The example used in this report is that of the selling price of homes sold in the last 100 days in Santa Rosa, CA. The homes that sold for low prices correspond to valleys in a terrain and the homes that sold for high prices correspond to peaks in the terrain.

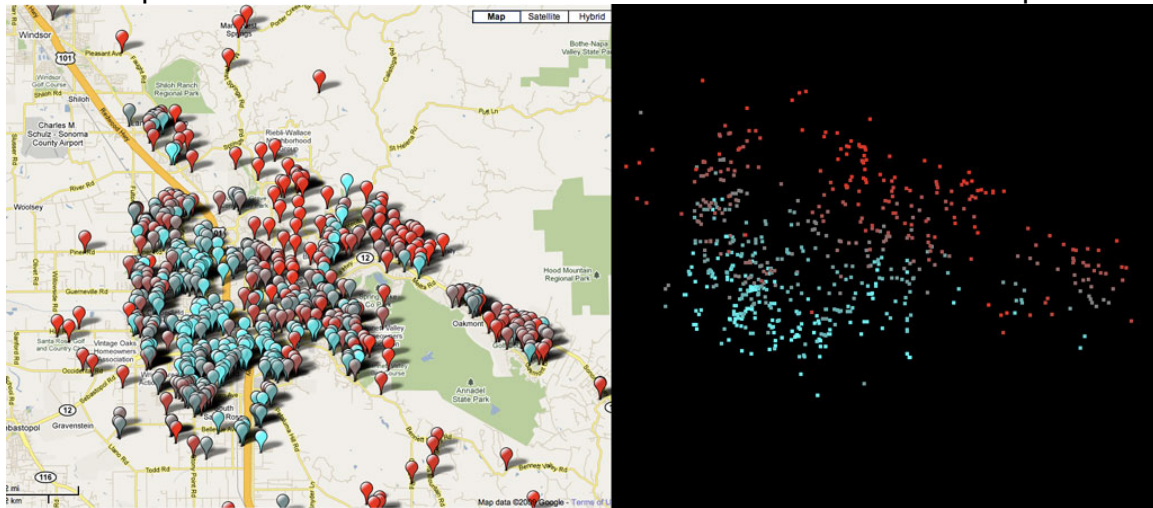


In the image above you can clearly see that the low end of the Santa Rosa market is west of the freeway, with the market median being towards the east and the high end surrounding and occupying much of the border.











## Generating the Surface

My implementation uses home sales data stored in a MySQL database, and uses map tiles served by Google Maps to generate the display. The primary problem is to generate a 3 dimensional surface from a set of samples where the geo-coordinates, longitude and latitude, corresponding to the x and y axes, remain fixed while the z value is allowed to “fit” the information field using the geo-coordinates of the sample data.

The image below shows the sample data first pinned onto a Google map and then as a point-cloud in 3D where the information field determines the z-position.

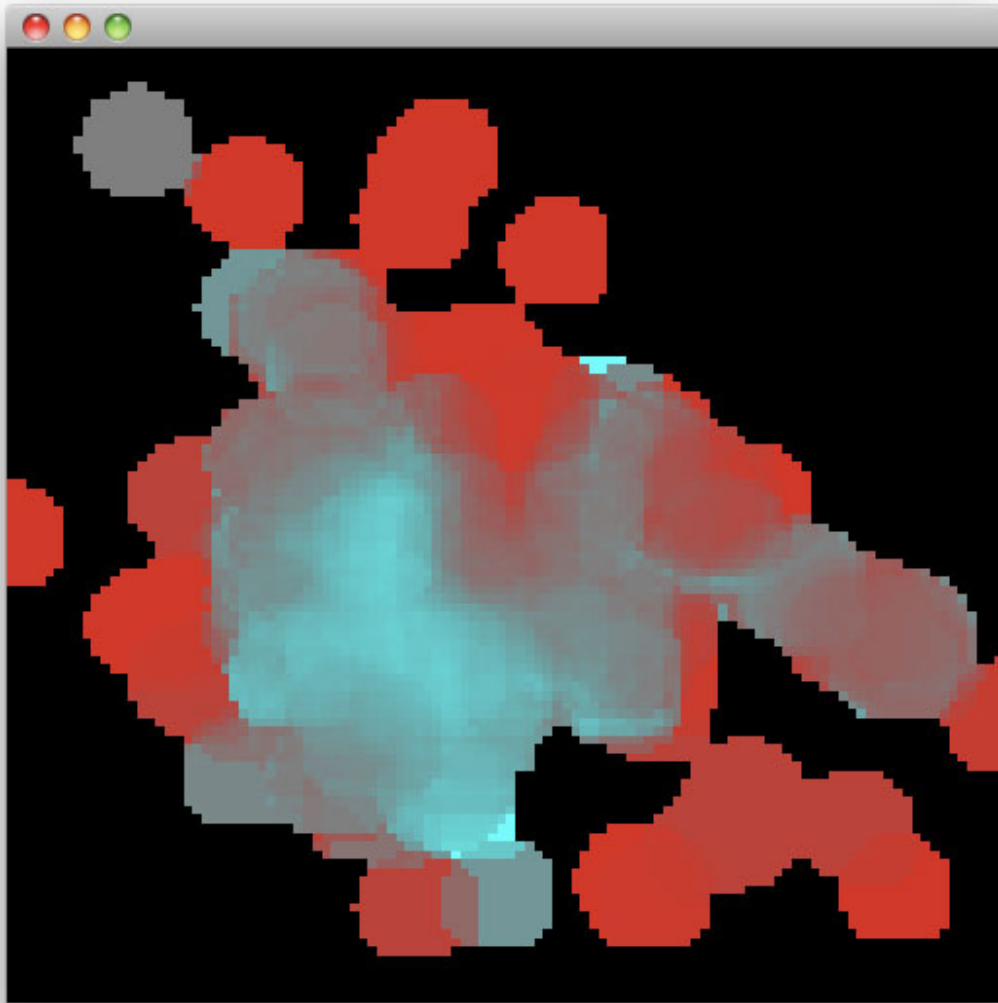


The coloring is also intended to help the visualization, with the cyan correlated with the low values and the red correlated with the high values.

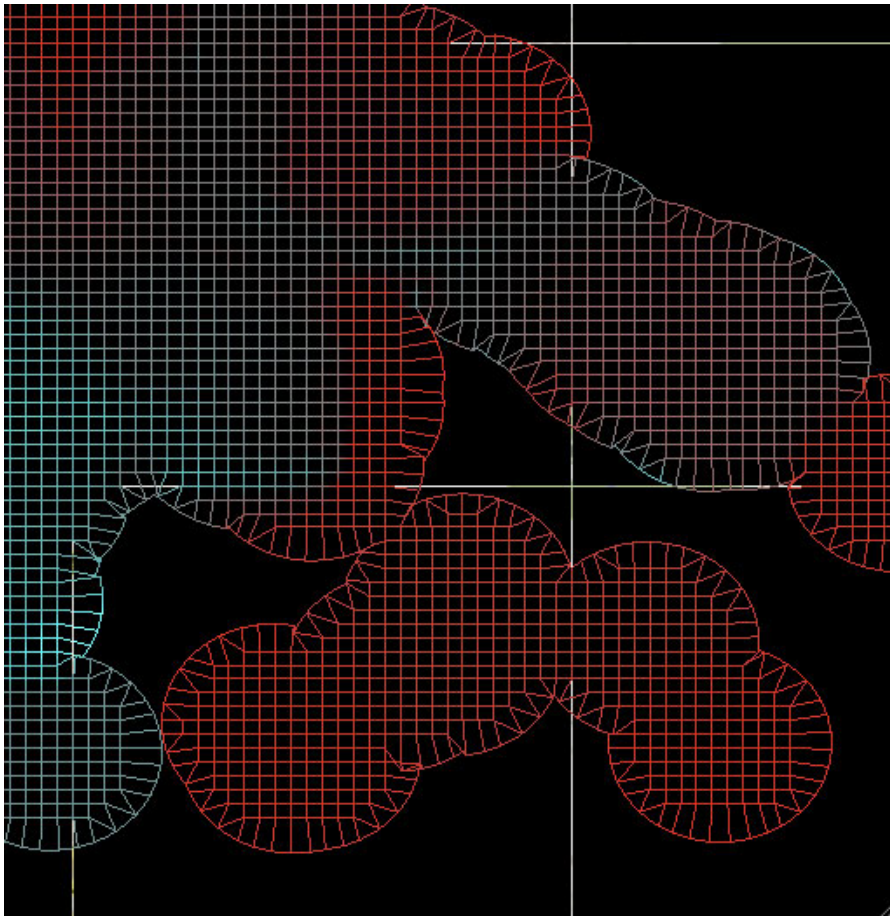
	\$582,500 - \$3,400,000	<p>It should be noted that the z value is not computed relative to the actual selling price of the property, but rather, according to the rank of the value. That is, the data is sorted and ranked so that the maximum value is twice as high as the median value. Without this measure, the most expensive properties would appear as skyscrapers towering over the landscape, as can be seen with this legend which shows the largest value being 10 times greater than the median value.</p>
	\$430,000 - \$582,500	
	\$370,000 - \$430,000	
	\$331,000 - \$370,000	
	\$300,000 - \$331,000	
	\$275,000 - \$300,000	
	\$250,000 - \$275,000	
	\$210,000 - \$250,000	
	\$160,000 - \$210,000	
	\$50,000 - \$160,000	

Next a technique for connecting a surface out of the sample data was in order. I chose to assign a strict radius of influence to each of sample points, of roughly a mile. This means that there’s a hard falloff at a distance of one mile from the property, but up until that fall-off the land is influenced just as much at all areas in the disk of influence. Overlapping disks are then averaged and stored in a 2

dimensional grid of pixels. This initial pixilated result can be seen below.



By creating a hard fall-off at one mile, two defects can be seen in the above image. One is that there is still a boxy cut-off in both the z-correlated coloring and the border of the surface. These two issues share a cause but are resolved in two different ways. For the borders, inspecting their 8 direct neighbors for a lack of surface identifies border pixels. If a pixel is determined to be a border pixel it is then pushed to the actual border by finding its nearest sample point and pushing it off in the opposite direction to the actual 1-mile edge of that sample's disk of influence. The effect this has on the pixel grid are shown in the next image.

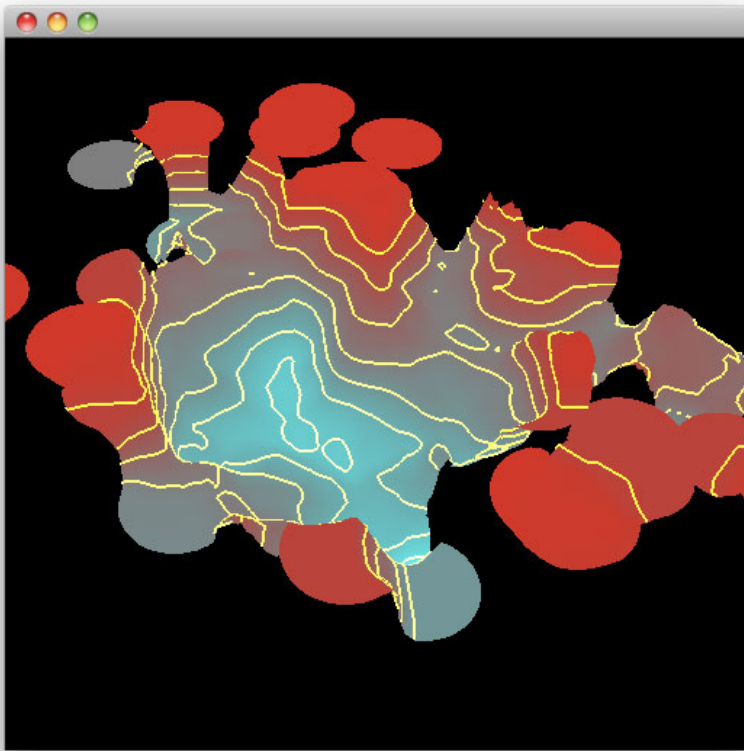


Next, to address the issue of the hard edges in the actual z-values, we apply a convolution step that again averages the values that lie within some fixed radius. This acts as a low pass filter blurring the colors, and smoothing the surface. This radius can be adjusted to the desired level of smoothing.

### **Contour Lines**

So far we have created an actual 3D surface tied together with a “thermal” coloring scheme, so that when viewed from above with an orthographic projection, the 3<sup>rd</sup> dimension is still visible in the coloring. Another great visual aid in addition to thermal coloring is to trace contour lines at a specified elevation. This contour line can show precisely where a given percentile division occurs. We can then create a hard divide between percentiles, showing for example a border between the properties sold below and above the median price.

Given a set of pixels with a height value the natural choice for computing these contours is by using the Marching Squares technique, and this is the technique I used in this implementation. The image below illustrates 9 contours “deciling” the surface with this technique.



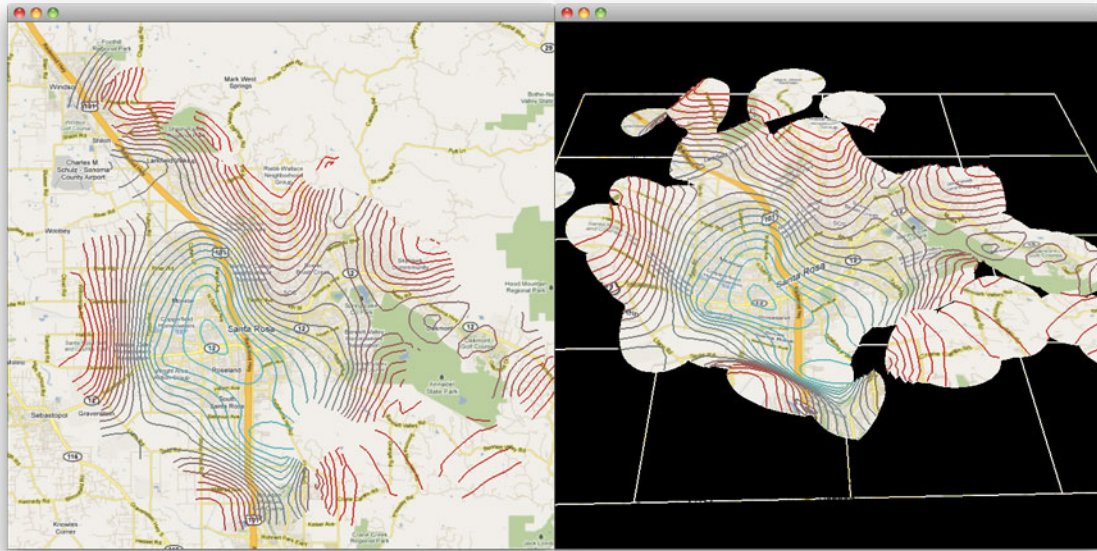
## Texture Mapping

Of course, the pretty height map visuals do not convey the information without the actual map image to tie the surface to streets and landmarks of the city being viewed. Thus, our final step in this implementation is to query the Google Map servers for corresponding tiles and applying it as a texture to the surface. We could blend the thermal colors into the map, or as with this implementation, we can remove the thermal coloring and leave just the contour lines with the thermal colors.

To get the visual to play perfectly with the Google tiles, the lon/lat coordinates needed to be projected into Google Maps tile/pixel space using the Mercator projection. To ensure a nice invisible seam in the tiles it makes sense then to create our own pixel grid in alignment with the Google tile/pixel grid. With the border pixel pushing in place, and the marching squares triangulation on our contour lines we are fine to use a much lower resolution pixel space than the Google pixels. This results in a height map resolution equivalent to 8 Google pixels, and a Google Map tile being equivalent to 256 Google pixels and 32 height map pixels. The 32 height map pixel alignment is visible in the grid image (4<sup>th</sup> image) included with this report (the white lines trace the Google tiles).



The next image shows this final result from an orthographic view where the contour lines imply the surface, and in a 3D view, but with more smoothing than was seen before.



## Future Work

I think that this project could greatly benefit from a “confidence mask” that takes into account the density of the samples within range of a pixel, and perhaps the distance from that pixel, since there can certainly be a dramatic shift in values within a mile. This confidence value could then be combined with the marching squares technique to compute a better border, or combined with an alpha channel to add a transparent fall-off effect.

Another important step to increasing the usefulness of this tool would be to display data at different zoom levels, panning out to the entire North Bay to see Marin County in red and Solano County in Cyan, or zooming into your favorite zip code to see a more local breakdown. Google Maps of course already provide the framework for this zooming that I can tie into, adjust the height map pixel grid accordingly. Perhaps wider zooms would call for a larger radius of influence, perhaps not. Handling the immense amount of sample points at a wide zoom or the sparse amount of sample points at a narrow zoom might also pose a problem, perhaps the data would just be queried far enough back in time to meet a density requirement, with wide zoom not going as far back in time as a narrow zoom.

Time series animations would be fun to do as well, we now have the basic framework in place for visualizing new construction, gentrification, and “white flight” for lack of a better term.

## **Conclusion**

Though real estate pricing is used specifically, this is a general way of depicting geographic trends of information in an intuitive and aesthetically pleasing way. As a project, it was fun to bring together both 2D and 3D ways of considering the implementation. Is it a 2D set of pixels or a surface in 3D space interpreted out of a point cloud? It's simultaneously both things, and that was an interesting insight for me as I thought about my techniques from both perspectives.