A simple approach to counting temporary shelters in Haiti

Robert Jones jones@craic.com

## Introduction

A number of areas of data analysis are dominated by sophisticated algorithms, intensive computation and, in some cases, limited access to raw data. Examples include the analysis of satellite imagery, feature extraction from video, protein structure analysis and language processing.

I am interested in how simple, approximate methods can be used to extend the application of these technologies. While simple approaches will clearly not match the accuracy and resolution of complex methods, they can, by nature of their simplicity, be implemented and deployed more easily and hence more widely.

Presented here is a very simple, low resolution approach to identifying and counting temporary shelters in Port-au-Price, Haiti. The approach uses publicly available images from Google Maps, as opposed to commercial imagery, combined with simple, non-proprietary software.

Chris Grundy at the London School of Hygiene and Tropical Medicine has worked with Medicins Sans Frontières on a more sophisticated approach to the same problem (http://www.newscientist.com/article/dn21846-satellite-images-help-doctors-count-people-from-space.html), with a test case involving refugee camps in Chad.

I am very interested in seeing how the approach described here might be developed further and applied in this and other areas of humanitarian need.

## Image Analysis

Satellite images of Port-au-Price, Haiti, from Google Maps have very distinctive features that represent temporary shelter for residents that lost their homes in the 2010 earthquake.

Numerous areas have a very high density of white, blue and sometimes rust colored small structures. Some of these areas are marked as internally displaced persons (IDP) camps, others may be small pockets of temporary and/or poor condition housing.

The small blue structures are almost certainly tents or shacks that are covered with blue plastic tarpaulins. These are particularly interesting as they are easily distinguished from other features.

I have been experimenting with simple approaches to quantifying the relative number of these structures in adjacent quadrants in these images. One application of this would be to monitor changes in the number and location of these structures over time. Those data could be very useful to organizations responsible for delivering aid and shelter to Haiti.

These four images illustrate the features. This area can be seen in Google Maps at http://goo.gl/maps/7ObKv

I have applied the approach to a 20 x 20 matrix of quadrants of 180m square. The location can be seen as the black square on Google Maps at http://goo.gl/eGfGZ.

The image on the next page shows that region and the table below shows the values computed in each of the quadrants.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 10565 | 16447 | 21212 | 4918 | 3712 | 3531 | 1743 | 3550 | 2980 | 1970 | 1879 | 2530 | 0 | 1416 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 870 | 10655 | 13168 | 343 | 1207 | 452 | 0 | 732 | 5812 | 3699 | 3480 | 528 | 366 | 2506 | 2979 | 3054 | 1086 | 0 | 335 |
| 2 | 2008 | 3945 | 0 | 1427 | 1478 | 0 | 0 | 1885 | 0 | 319 | 2126 | 0 | 0 | 0 | 0 | 0 | 408 | 0 | 0 | 1418 |
| 3 | 1432 | 3816 | 1111 | 872 | 4646 | 1633 | 0 | 0 | 0 | 1967 | 1172 | 0 | 0 | 0 | 657 | 0 | 1401 | 846 | 1012 | 154 |
| 4 | 2350 | 3467 | 1763 | 1947 | 969 | 4549 | 510 | 645 | 649 | 2571 | 0 | 530 | 0 | 534 | 928 | 1411 | 1004 | 255 | 0 | 498 |
| 5 | 0 | 1093 | 699 | 1586 | 944 | 964 | 2577 | 2186 | 1284 | 2799 | 321 | 623 | 0 | 4008 | 515 | 0 | 3225 | 0 | 0 | 749 |
| 6 | 1130 | 4011 | 2518 | 0 | 4054 | 1371 | 6019 | 3578 | 1842 | 5022 | 0 | 1563 | 1934 | 2164 | 543 | 1746 | 2794 | 0 | 0 | 0 |
| 7 | 3132 | 2366 | 3322 | 510 | 1163 | 7039 | 4113 | 3723 | 8880 | 4246 | 7364 | 3343 | 1040 | 1062 | 2168 | 2527 | 1820 | 523 | 0 | 182 |
| 8 | 2939 | 56 | 982 | 2185 | 3949 | 6349 | 5151 | 3894 | 2052 | 3230 | 1543 | 4045 | 3325 | 1696 | 1994 | 2426 | 2252 | 2631 | 660 | 0 |
| 9 | 546 | 0 | 1037 | 5067 | 3133 | 4542 | 4153 | 5568 | 2376 | 1475 | 1102 | 2675 | 1216 | 2852 | 0 | 0 | 13049 | 11215 | 1899 | 2341 |
| 10 | 920 | 6796 | 0 | 1829 | 1673 | 1191 | 1011 | 3095 | 4086 | 2771 | 3835 | 1806 | 0 | 1267 | 0 | 1399 | 9259 | 9238 | 4525 | 1023 |
| 11 | 0 | 838 | 7913 | 2736 | 0 | 434 | 0 | 871 | 1702 | 4297 | 1543 | 2474 | 8058 | 3285 | 860 | 2321 | 1812 | 3884 | 773 | 0 |
| 12 | 1550 | 3375 | 10169 | 1201 | 0 | 1054 | 0 | 1323 | 7904 | 4946 | 4858 | 7636 | 8935 | 4753 | 1242 | 1584 | 3105 | 4102 | 2878 | 1522 |
| 13 | 0 | 1524 | 10776 | 2151 | 1024 | 743 | 948 | 887 | 1061 | 617 | 882 | 1684 | 1396 | 3247 | 1150 | 1342 | 2681 | 0 | 2734 | 1884 |
| 14 | 4202 | 2152 | 3561 | 0 | 2338 | 473 | 0 | 0 | 0 | 867 | 561 | 320 | 8777 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 9004 | 4240 | 749 | 481 | 2481 | 838 | 0 | 0 | 0 | 0 | 0 | 128 | 0 | 1367 | 2073 | 0 | 0 | 0 | 0 | 0 |
| 16 | 6059 | 2560 | 0 | 663 | 519 | 3134 | 574 | 1149 | 397 | 520 | 0 | 324 | 0 | 3134 | 4130 | 485 | 0 | 0 | 0 | 0 |
| 17 | 2582 | 466 | 893 | 0 | 0 | 844 | 340 | 578 | 0 | 905 | 0 | 0 | 1045 | 648 | 0 | 0 | 0 | 2245 | 688 | 0 |
| 18 | 4056 | 1149 | 471 | 1219 | 31 | 0 | 316 | 0 | 231 | 749 | 414 | 1036 | 1385 | 951 | 0 | 1216 | 0 | 3953 | 1792 | 0 |
| 19 | 1051 | 1101 | 312 | 0 | 0 | 0 | 460 | 337 | 0 | 511 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 557 | 0 |

The overview image is too low resolution to make out the detail that is present in the quadrants, but the general locations of the camps/slums should be apparent. The high values in the table appear to correlate well with those areas. Note that all computed areas are relative and would require some calibration to produce absolute numbers of dwellings.

The steps involved in this approach are:

1: Fetch a defined quadrant from Google Maps
2: Split the color image into separate RGB channels
3: Subtract the maximum of the Red and Green planes from the Blue plane
4: Apply a threshold value that masks out low intensity blue pixels
5: Dilate the thresholded image slightly to reduce noise
6: Find contours in the image
7: Compute the area for each contour and ignore those over a cutoff (which may represent large buildings that happen to have blue roofs)
8: Sum the areas of all contours in the quadrant

An example of the thresholded image and the contours overlaid on the original is shown here:

The difference/threshold approach appears to be effective in find the blue features in the image on the left side. The use of contours to eliminate large structures works well in other example images.

However, in this example there are a number of blue features in the lower right corner that are very close to each other and contour detection has placed these into a single feature with an area greater than the cutoff. Clearly the approach is preliminary.

The approach does not detect any specific feature. A more thorough and accurate approach would attempt to recognize the rectangular shapes of all roofs and filter those based on, perhaps, size and density.

However, this simple approach may have utility in some applications. First off, it is very simple making it easy to modify without expert knowledge of image processing. It uses open source software (OpenCV and Python scripts) and images from Google maps, which makes it available to a wider audience than proprietary software such as that from ESRI. It also uses relatively low resolution images which, again, may be important for widespread use by those without access to high resolution imagery.

All the code is freely distributed and can be found at https://github.com/craic/count_shelters.