Towards Quantifying the Amount of Uncollected Garbage through Image Analysis

Susheel Suresh, Tarun Sharma, Prashanth T.K., Subramaniam V, Dinkar Sitaram & Nirupama M

PESIT

Centre for Cloud Computing and Big Data

Problem

One of the most common problems faced by metropolitan and emerging cities, is the problem of garbage collection and disposal. Even in the city Bangalore (India), the local municipal committee (BBMP) estimates that 3,500 tonnes of garbage is produced by the city everyday. Despite considerable efforts, nearly 20 per cent of the waste still remains to be picked or is picked irregularly, giving the city a despicable look.

Problems related to quantification of garbage from a computer vision perspective.

1. No existing image dataset for garbage analysis.
2. No automatic detection and reconstruction pipeline.

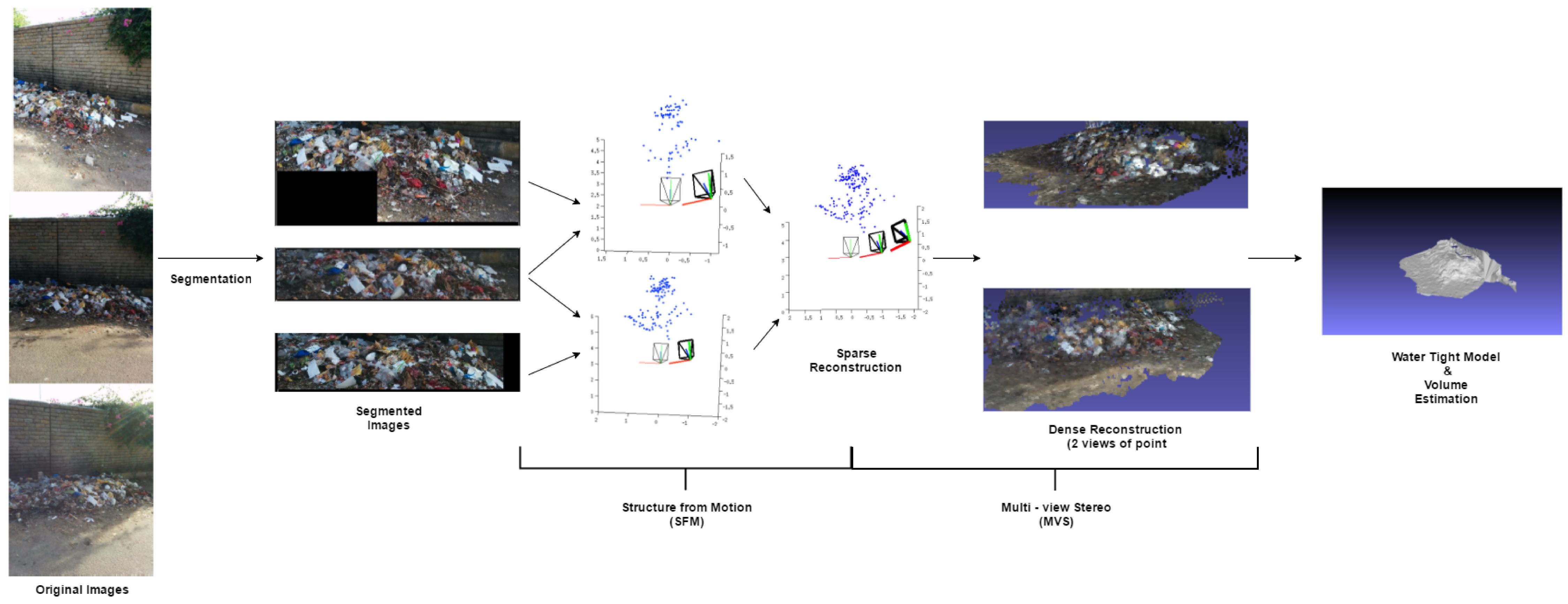
Contributions

We address the problem using a two step approach. In the first step, we build a mobile application that allows citizens to capture images of garbage and upload them to a server. In the second step, back-end performs analysis on these images to estimate the amount of garbage. Using our novel pipeline, experiments indicate that with 8 different perspectives, we are able to achieve an accuracy of about 85 % for estimating the volume.

Segmentation

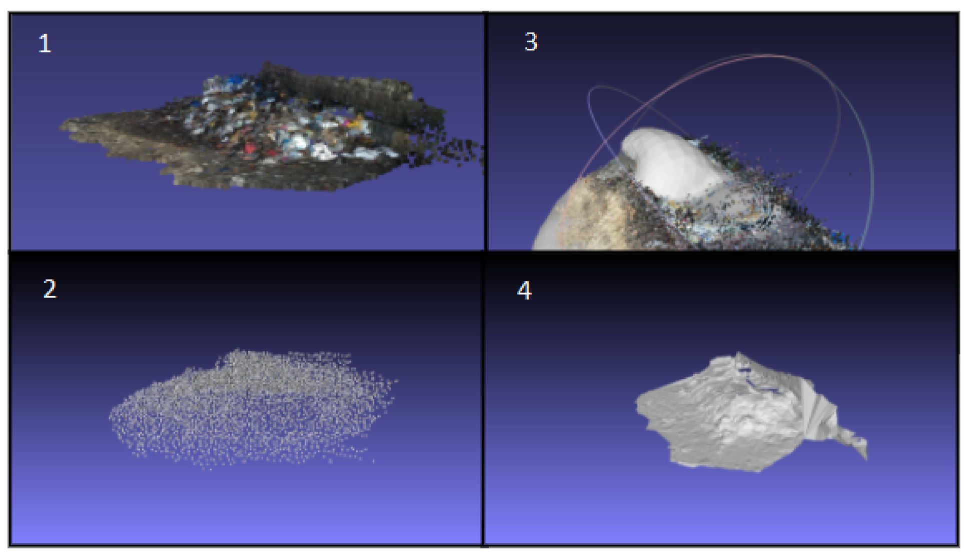
The first step in the pipeline is to segment out the garbage from the background. We compared three methods of segmenting out the garbage, sliding window with edge thresholding, sliding window with a fully connected neural network, and a bounding box approach using deep learning. In the first method, we assume that the number of edges in an area with garbage will be beyond a threshold when compared to the background. This works well against walls but fails to remove trees. For the second method, we train a classifier using neural networks to classify each window as garbage or non-garbage. This addresses the shortcomings of method 1 but does not give us a proper rectangular shape, and hence affects the next stages. The third method involves a fine tuned AlexNet with its last layer modified, trained to regress to the coordinates of the bounding box.

Pipeline



The input to the system is a set of multi-view images of garbage dumps. The first stage detects and segments the garbage from the scene. It is followed up with structure from motion and multi-view stereo techniques to produce dense reconstructions. Finally after generating water tight 3D models, volume is estimated.

3D Reconstruction



.

We use a typical incremental SFM system, where two view reconstructions are first estimated upon successful feature matching between two images, 3D models are then reconstructed by initializing from good two-view reconstructions, then repeatedly adding matched images, triangulating feature matches, and bundle-adjusting the structure and motion. Using a set of 8 multi-view images for a scene, we gather the camera properties, perform feature matching using SiftGPU, then apply the incremental Structure from Motion algorithm to generate sparse reconstructions. For the final leg, we use Furukawa and Ponce's PMVS algorithm to produce dense reconstructions.

Surface Reconstruction

We take the point cloud from Stage 2 obtained from multi-view stereo and 3D SFM. The aim of this stage is to estimate the volume of garbage in the scene, which is done by estimating the volume of the point cloud. We first provided a surface for the scene-object by using different methods of surface reconstruction. The Ball-Pivot point method gave the most efficient results with regard to computation time and resources. We then make the mesh (reconstructed surface) manifold by closing all holes and making the model completely watertight. We then estimate volume of the watertight model using an experimentally determined scale factor empirically.

Results

The volume of each garbage dump was measured while the image dataset was being created. When we were creating the image dataset, we manually measured the approximate enclosure boundary volume (length, breadth and height) of each garbage dump.

Location	Measured actual volume (m^3)	Computed volume from pipeline (m^3)	Percent Error (%)	Total Execution Time (s) Segmentation+ Reconstruction+Volume estimation
Domlur	1.198	0.999	16.61	$2.2 + 243 + 1.5 = 246.7$
BTM Layout	3.442	3.029	11.99	$2.4 + 257 + 1.7 = 261.1$
Marathahalli	0.855	0.982	14.85	$2.1 + 248 + 1.4 = 251.5$
Nayandanahalli	1.402	1.119	20.18	$2.7 + 261 + 1.7 = 265.4$

Comparative Study - A high error of 24.30% was achieved when edge thresholding and poisson reconstruction were used. It dropped to 22.13% when the same was accompanied with ball-pivot instead of poisson reconstruction. We attained lower error percentage values of 19.32 and 16.19 when we used sliding window with neural networks in our first stage. The former was followed up with poisson and later with ball-pivot. Using the bounding box approach with CNN features in the initial stage, we arrived at the lowest error percentages of 14.61 and 12.43, with poisson and ball pivot respectively.