Motivation
How can we visualize places that no longer exist or never existed?
Piazza San Marco with the Basilica, Venice. 1730
The Crystal Palace. 1851.
Bedroom in Arles. 1888.
Background
Perspective in Art

- Renaissance and Roman artists experimented with perspective to create more realistic depictions.
- Romans used a technique that resulted in several layers of the image, each with their own set of vanishing points.
  - Difficult to reconstruct
- Renaissance artists used a more true technique.
  - Easier to reconstruct
Interior images often have one point perspective. Exterior images often have more.
Steps

- Vanishing Point Detection
- Ground Detection/Plane Detection
- Object Segmentation
- Depth Calculation
- Normal Calculation
- 2.5D Reconstruction
Techniques
Vanishing Point Detection
Vanishing Points

- A vanishing point is a point in the picture plane that is the intersection of the projections (or drawings) of a set of parallel lines in space on to the picture plane.
- The shorter the blue line the further away on the z-axis the pixel is.
- \( Z(x, y) \) is proportional to \( \frac{1}{d} \).
Edge Detection

- Blur image using Bilateral Filtering
- Use Sobel Operator on image to calculate gradients
- Normalize gradients and remove pixels where gradients have too small magnitude
- Hough Line
Vanishing Point and Converging Lines

- Calculate intersections of lines
- Compute representative point
- Remove points that are a certain sigma away.
- Regress for a certain amount of iterations (we used 100)
- Vanishing point is that representative point
- Calculate converging lines using the dot product
Vanishing Point and Converging Lines (cont.)

- Currently only supports one vanishing point
- How can we support two or three?
  - Perform clustering on intersection points
    - Results are okay but have a lot of error
    - We will have to perform some sort of regression!
- What is the input to our regressive technique?
  - The centers of the clusters found (using k means for example)
  - Choose a large K, then determine validity of center by how many lines converge on it
Regression Technique

- Cluster Center as input
- Vanishing Point as output
- Calculate 2D Bivariate Distribution
  - Remove points that exist outside of probability
- Use “learning rate” to mimic certainty as number of iterations increase
- Increase minimum acceptance probability as iterations increase
Does this work?
Mostly.
Regression Technique (cont.)

- Given good enough seed points, the regression technique will converge to the correct vanishing points.
- However, if vanishing points are too close, KMeans++ will fail.
- One technique we could use is Statistical Outlier Removal.
Vanishing Points as of Today

- Very dependent on input lines
- Works very well with one point perspective
  - $D(\text{VP}', \text{VP}) < 2\%$ on average
  - Works well enough to use as input to ground detection
- Works well for multi point perspective but not consistent enough
- For now, restrict paintings to paintings with either one point perspective or paintings that can be approximated with one point perspective
Ground Detection
Ground Detection

Assumptions:

- Assume our inputs are drawn/taken in one point perspective and is tunnel-like.
- Ground and walls are perfectly perpendicular.
- The objects we want to extract are perpendicular to the ground or the walls.
- The objects are not overlapping in the image

Ground is parallel to xz plane
Walls are parallel to yz plane
Reference to depth calculation
Ground Detection cont.

All the straight lines that pointing to the vanishing point are possibly parallel to z axis. (the blue lines showing on the right)

Vertical lines (possibly parallel to y axis).
Horizontal Lines (possibly parallel to x axis)

There are more lines on the ground that parallel to x axis since ground is parallel to xz-plane.

The ratios of vertical and horizontal lines determines the normal of the plane.
Ground Detection cont.

Draw a line to separate the end-points of the vertical lines and horizontal lines.

The line will pass through the vanishing point.

\[
\Theta = \max \left( \frac{C_1 P_a}{P_a + P_b} + \frac{C_2 P_b}{P_a + P_b} \right)
\]

On the two lower quarters of the image.
Non-symmetric photograph
Segmentation
Segmentation and Anchor Point

- Goal of segmentation is to separate the objects from the floor and sky/background

- Input:
  - Source Image

- Steps:
  - Blur using Bilateral filter
  - Threshold using adaptive threshold
  - Remove noise using dilation/erosion
  - Find contours
  - Ignore contours that have contour area below/above a threshold
  - Use bottom right point of contour bounding rectangle as anchor point

- Output:
  - Vector of contour points
  - Vector of anchor points
  - Matrix with pixel to anchor point mapping
Depth Calculation
Depth Calculation

- In order to construct the 2.5D image the segmented objects need to be placed on the z-axis and we can calculate these z values using the depth function.

  **Input:**
  - Vanishing point (assuming 1 VP)
  - Vector of anchor points (AP)
  - Vector of 2 points for road/floor lines

  **Steps:**
  - Calculate if anchor point is on the floor or walls
  - Floor:
    - Get distance between VP and AP y-values
Depth Calculation (cont.)

- Walls:
  - Calculate if AP is to the left/right of VP using x-values. Use this to see if we should use the left floor line or right floor line
  - Calculate y-value using AP x-value and equation of appropriate floor line
  - Get distance between VP and AP y-values

- Output:
  - Depth values for each anchor point
  - Using the mapping of pixels and their anchor points we give each pixel a depth value for the 2.5D construction
Segmentation and Depth Future Work

- **Limitations:**
  - Segmentation is dependant on parameters so the parameters need to be optimized to get good segmentation
  - Only works for 1 vanishing point

- **Future Work:**
  - Use cross ratio to get better estimates for the z-values
  - Optimize parameters used for object segmentation
  - Determine z values given 1, 2 or 3 vanishing points

The cross-ratio \((ABCD)\) of four collinear points \(A, B, C, D\) is defined as the "double ratio":

\[
(ABCD) = \frac{CA}{CB} : \frac{DA}{DB},
\]
Potential Applications
Applications

- Reconstruction of non-existent scenes
  - Games
  - Solving Crime
  - Ancient Wonders
- Road detection for self-driving cars
- Proximity detection
- Detecting layers in images
Questions?
Don’t ask us if it works!