3D RECONSTRUCTION OF A BUILDING FROM SINGLE IMAGE

Yawen LIU
Department of Information Engineering
Wuhan Technology University of Surveying and Mapping
39 Luoyu Road, Wuhan, P. R. China (430079)
e-mail: yuluyk@263.net
Zuxun ZHANG
Department of Information Engineering
Wuhan Technology University of Surveying and Mapping
Jianping ZHANG
Department of Information Engineering
Wuhan Technology University of Surveying and Mapping

Keywords: orientation parameters the distance ratios 3D reconstruction

Abstract
In architectural photogrammetry, the buildings are designed with a few basic shapes with parallel lines. So, in this paper a new approach to 3D-reconstruction of building from single images is presented. The method requires three sets of parallel lines. In the first step, three couples of parallel straight lines are used to determine the rotation parameters, interior orientation parameters and translate parameters. In the second step, the ratios of the distance from the projection center to the corner points of the parallelogram are computed using only parallelity information. With the result of two steps, 3D coordinates of building can be computed. The method was based on stronger theory of mathematics. So it's robust, accurate.

1 Introduction
The 3D reconstruction of buildings has been an active research topic in computer vision as well as in digital photogrammetry in recent years. Three dimensional building models are increasingly necessary for urban planning, tourism. Manual 3D processing of images is very time consuming. Therefore, speeding up this process by automatic or semi-automatic procedures has become a necessity. There are a lot of systems that work solely with monocular images. These systems exploit shadows either to infer the third dimension or to verify the generated hypothesis. Other systems use widely stereo images, the determination of the third dimension by epipolar matching of different features extracted from both images.

The approach presented in this paper, works solely with monocular images. The buildings are assumed to be rectangular or rectilinear flat roofs. The procedure consists of two steps. In the first step, the interior and exterior orientation parameters can be determined with parallel lines information and two object control points. In the second step, the ratios of distance from the projection center to the corner points of the parallelogram are computed using only parallelity information. The method makes full use of linear features and constrains(co-planar, parallel, vertical) to determine geometric and mathematical relations.

2 Determination of interior and exterior orientation parameters
From figure 2-1, we can see that the perspective center, straight line in 3-D object space and its projection on image plane can form a plane. The co-planar constrain is the foundation of this new algorithm.

![Figure 2-1 Interpretation plane and normal vectors](image)

In order to have a good understanding of this, we take the following symbols: L: straight line in the object space P: any point on the image plane S: perspective center of image N: the normal vector of plane defined by S and L P: any point of P on the image plane o: the center of the image plane c-x-y: image plane coordinate system S-x-y-z: image space coordinate system O-X-Y-Z: object space coordinate system

Notation: vectors are printed in bold
Suppose $\vec{x}_{0y}$ is the coordinate of the principle point in image plane coordinate system o-x-y, $(x, y, f)$ is the coordinate of point $r$ in the image space coordinate system S-x-y-z. The direction vector of $\vec{l}$ in S-x-y as

$$\vec{l} = \begin{bmatrix} v_x \\ v_y \\ 0 \end{bmatrix}$$

where $v_x$, $v_y$, and $0$ are three coordinates of $\vec{l}$ on the x-, y-, and z-axes respectively.

The direction vector of $\vec{S}p$ as

$$\vec{S}p = \begin{bmatrix} x_i - x_0 \\ y_i - y_0 \\ -f \end{bmatrix}$$

We can obtain the normal vector of plane defined by S and L in S-x-y-z through the cross-product of $\vec{l}$ and $\vec{S}p$

$$\vec{n} = \begin{bmatrix} -v_y \cdot f \\ v_x \cdot f \\ d_i - v_x \cdot y_i + v_y \cdot x_i \end{bmatrix}$$

where $d_i = v_x \cdot y_i - v_y \cdot x_i$

Suppose $v_x, v_y$ and $v_z$ are three coordinates of L on the x-, y-, and z-axes in the image space coordinate system S-x-y-z respectively.

$$\vec{L} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$$
The plane defined by perspective center of image S and straight line I is called interpretation plane. According to coplanar coordinate of perspective geometry, the straight line L in the object space is on the plane defined by the straight line I (projection of L) and Sp. So, L is perpendicular to the normal vector of plane defined by S and I in S-xyz, the dot product of L and n is equal to zero.

\[-v_y \cdot x + v_z \cdot y + v_x \cdot z = 0 \quad (5)\]

Equation (5) may be expressed by

\[-v_y \cdot \alpha + v_z \cdot \beta + \frac{d_y \cdot v_x - v_x \cdot d_z + v_z \cdot d_y}{f} = 0 \quad (6)\]

where

\[\alpha = \frac{v_x}{v_z} \quad \beta = \frac{v_y}{v_z} \quad (7)\]

When L is parallel to the X-axis in the object coordinate system, we can transform vector L from O-XYZ to S-xyz with the rotation matrix R and scale \(a\).

\[
\begin{bmatrix}
  x' \\
  y' \\
  z'
\end{bmatrix} = R \begin{bmatrix}
  x \\
  y \\
  z
\end{bmatrix} + \begin{bmatrix}
  a \\
  b \\
  c
\end{bmatrix}
\]

where \(a, b, c\) are three coordinates of L on the X-, Y- and Z-axes in the object coordinate system respectively.

Substituting equation (6) into equation (6),

\[
\begin{align*}
\alpha &= \frac{a_1}{a_3} \\
\beta &= \frac{a_2}{a_3}
\end{align*}
\]

where \(R = (a_1, b_1, c_1)^T\), \(R = (a_2, b_2, c_2)^T\) is the rotation matrix defining the camera orientation, \(a_1, b_1, c_1\) are the direction cosines or the elements of the rotation matrix \(R\).

Suppose Hand I2 in the image plane are the projections of two straight lines parallel to the X axis in the object coordinate system. Equation (6) can be given by another expression of the following form respectively,

\[
\begin{align*}
-v_y \cdot \alpha + v_z \cdot \beta + k \cdot d_y - k \cdot d_z + v_z \cdot y_0 + k \cdot v_y \cdot x_0 &= 0 \\
-v_y \cdot \alpha + v_z \cdot \beta + k \cdot d_y - k \cdot d_z + v_z \cdot y_0 + k \cdot v_y \cdot x_0 &= 0
\end{align*}
\]

where \(k = \frac{1}{f}\)

According to equation (10), we can obtain

\[
\begin{align*}
\alpha_1 &= k \cdot (h_1 + x_0) \\
\beta_1 &= k \cdot (g_1 + y_0)
\end{align*}
\]

\[
\begin{align*}
h_1 &= \frac{v_x \cdot d_z - v_y \cdot d_z}{v_z \cdot v_z} \\
g_1 &= \frac{v_x \cdot d_z - v_y \cdot d_z}{v_z \cdot v_z}
\end{align*}
\]

Similarly, when Li is parallel to the Y-axis in the object coordinate system or Li is parallel to the Z-axis in the object coordinate system we can get

\[
\begin{align*}
\alpha_2 &= k \cdot (h_2 + x_0) \\
\beta_2 &= k \cdot (g_2 + y_0) \\
\alpha_3 &= k \cdot (h_3 + x_0) \\
\beta_3 &= k \cdot (g_3 + y_0)
\end{align*}
\]

where \(h_0, g_0, h_i, g_i\) have meaning similar to \(h, g\), respectively.

Considering orthogonal constraints of the nine elements in the rotation matrix \(R\), we can get the following equations

\[
\begin{align*}
h \cdot h + g \cdot g &= (h_1 + h_0) \cdot x_0 + (g_1 + g_0) \cdot y_0 + z_0^2 + y_0^2 + f^2 = 0 \\
h \cdot h + g \cdot g &= (h_1 + h_0) \cdot x_0 + (g_1 + g_0) \cdot y_0 + x_0^2 + y_0^2 + f^2 = 0
\end{align*}
\]

where \(h_0, g_0, h_i, g_i\) have meaning similar to \(h, g\), respectively.

If we know the interior orientation parameters \(x_0, y_0, f\), for camera, considering the equation (9),(12),(14),(15), we can get

\[
\begin{align*}
a_1 &= \frac{1}{\sqrt{\alpha_1^2 + \beta_1^2 - 1}} \\
c_3 &= \frac{1}{\sqrt{\alpha_1^2 + \beta_1^2 - 1}} \\
b_3 &= \frac{1}{\sqrt{\alpha_2^2 + \beta_2^2 - 1}}
\end{align*}
\]

Substituting equation (17) into equation (9) and (14), we can obtain nine elements of \(R\).

According to collinear equations

\[
\begin{align*}
x - x_0 &= -a_1 \cdot (X - X_0) + b_1 \cdot (Y - Y_0) + c_1 \cdot (Z - Z_0) \\
y - y_0 &= -a_1 \cdot (X - X_0) + b_1 \cdot (Y - Y_0) + c_1 \cdot (Z - Z_0)
\end{align*}
\]

If knowing the rotation matrix \(R\), the camera interior orientation parameter \(x_0, y_0, f\) find two control points, according to Taylor's theorem, Eqs(18) may be expressed in linear form as

\[
\begin{align*}
x &= (x_0 + \frac{\partial x}{\partial X_0} \cdot dX_0 + \frac{\partial x}{\partial Y_0} \cdot dY_0 + \frac{\partial x}{\partial Z_0} \cdot dZ_0) \\
y &= (y_0 + \frac{\partial y}{\partial X_0} \cdot dX_0 + \frac{\partial y}{\partial Y_0} \cdot dY_0 + \frac{\partial y}{\partial Z_0} \cdot dZ_0)
\end{align*}
\]

where \((x_0, y_0)\) are approximate of function \(X_0, Y_0, Z_0\) are corrections of \(X_0, Y_0, Z_0\) respectively. \(X_0, Y_0, Z_0\) are unknown parameters. \(x, y\) are observations. So, we can get the adjustment model and the least-square solution to equation (19).

3 Computation of the distance ratios

These lines in the image are assumed to correspond to lines in object space that are coplanar and parallel(see figure3-1). The image coordinates of the 4 corner points \((l, j, k, l)\) of the parallelogram are arrived by measure manually or line intersection after edge detection. In the camera system, the image point can be expressed in vector built from the image coordinates \((x, y)\) and the focal length \(f\). This is the vector \((x, y, f)\).
Figure 3-1: parallelogram \(i,j,k\) and volume of a tetrahedron. The volumes of the 4 tetrahedrons (\(V_{i,j,k}, V_{i-k,j}, V_{i-j,k}, V_{i-j-k}\)) can be computed from the 3 vectors defined by the image coordinates, the focal length and the distances from the projection center to the 3 corner points. Volume of the tetrahedron (\(V_{i-j-k}\))

\[
V_{i-j-k} = \frac{d i j k}{6|i| j |k|} \tag{20}
\]

where

\[
|i| \quad |j| \quad |k| = \det(i, j, k)
\]

\[
x_i = i = (x_i, y_i, -f)
\]

\[|i|, |j|, |k| \] express the size of the vectors \(i, j, k\).

\[d_{i,j,k} \] express the distance between \(i, j, k\).

Because of the parallelity, a diagonal of the parallelogram always splits its area in 2 equal parts. And thus, the areas of the 4 triangles (\(\Delta A_{0} A_{1} A_{2} A_{3}\)) in the plane of the parallelogram are all equal. As the same holds for the distance \(h\) of the projection center to the plane of the parallelogram, it follows that the volumes of the 4 tetrahedrons are identical.

\[
\frac{dd_{i,j,k}}{|i| |j| |k|} = \frac{dd_{j,i,k}}{|j| |i| |k|} = \frac{dd_{k,i,j}}{|k| |i| |j|} = \frac{dd_{i,k,j}}{|i| |k| |j|} \tag{21}
\]

The distance ratios \(C_{i,j,k}\) can be written as follows

\[
C_{i,j,k} = \frac{d_{i,j,k}}{d_{i,j,k}} = \frac{k, i, j}{j, k, l} \tag{22}
\]

\[
C_{i,j,k} = \frac{i, j, k}{j, k, l}
\]

\[
C_{i,j,k} = \frac{i, j, k}{j, k, l}
\]

4 model coordinates and 3D coordinates

Model coordinates \(X^{\text{mod}}\) can be written as a function of one distance \(d_i\) that defines the scale of the model

\[
X^{\text{mod}} = \frac{d_i}{d_i} \cdot X(= \frac{d_i}{d_i}) = \lambda i x_i
\]

\[
X^{\text{mod}} = \lambda C_{i,j,k} X
\]

From the first step, interior and exterior orientation parameters are known. So, object coordinates can be computed in following form

\[
\Delta R X = X - X^0
\]

\[
\Delta R C (X) = X - X^0
\]

where \(\lambda\), scale factor

\(R\) rotation matrix

\(C_{ij}\) distance ratio projection center-points \(i\) and \(j\) from step 2.

Table 5-1 The simulated orientation parameters

<table>
<thead>
<tr>
<th>Interior orientation parameters</th>
<th>(x_0)(pixel)</th>
<th>(y_0)(pixel)</th>
<th>(F)(pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\phi)</td>
<td>-274.50</td>
<td>-236.96</td>
<td>2604.76</td>
</tr>
<tr>
<td>(\omega)</td>
<td>0.5642</td>
<td>0.9172</td>
<td></td>
</tr>
<tr>
<td>(\kappa)</td>
<td>57910.90</td>
<td>-47492.94</td>
<td>54088.57</td>
</tr>
</tbody>
</table>

Table 5-2 The simulated control points

<table>
<thead>
<tr>
<th>Number</th>
<th>(X_1)(mm)</th>
<th>(Y_1)(mm)</th>
<th>(Z_1)(mm)</th>
<th>(u_1)(pixel)</th>
<th>(v_1)(pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>492.26</td>
<td>-40.85</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>490.25</td>
<td>-41.89</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>-490.57</td>
<td>-39.60</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>-492.54</td>
<td>-38.58</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>-490.19</td>
<td>-38.51</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>-490.51</td>
<td>-37.22</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>-488.54</td>
<td>-38.25</td>
</tr>
</tbody>
</table>
Table 5-3 computation result

<table>
<thead>
<tr>
<th>Interior orientation parameters</th>
<th>$x_0$(pixel)</th>
<th>$y_0$(pixel)</th>
<th>$f$(pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-274.50$</td>
<td>$-236.98$</td>
<td>$2604.76$</td>
<td></td>
</tr>
<tr>
<td>Exterior orientation parameters</td>
<td>$\phi$</td>
<td>$\omega$</td>
<td>$K$</td>
</tr>
<tr>
<td>$-0.6810$</td>
<td>$0.5642$</td>
<td>$0.9172$</td>
<td></td>
</tr>
<tr>
<td>$X_s$(mm)</td>
<td>$Y_s$(mm)</td>
<td>$Z_s$(mm)</td>
<td></td>
</tr>
<tr>
<td>$57910.90$</td>
<td>$-47492.94$</td>
<td>$54088.57$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cj</th>
<th>$x$(pixel)</th>
<th>$y$(pixel)</th>
<th>$X^m$</th>
<th>$Y^m$</th>
<th>$Z^m$</th>
<th>$X$</th>
<th>$Y$</th>
<th>$Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0007</td>
<td>-217.71</td>
<td>126.13</td>
<td>-7673.37</td>
<td>6912.71</td>
<td>-91805.01</td>
<td>-2.53x10^4</td>
<td>-2.08x10^4</td>
</tr>
<tr>
<td>2</td>
<td>1.0001</td>
<td>-215.74</td>
<td>195.10</td>
<td>-7599.46</td>
<td>6872.26</td>
<td>-91751.15</td>
<td>100</td>
<td>2.41x10^4</td>
</tr>
<tr>
<td>3</td>
<td>0.9994</td>
<td>-216.07</td>
<td>197.39</td>
<td>-7605.50</td>
<td>6947.91</td>
<td>-91686.03</td>
<td>100</td>
<td>-2.83x10^4</td>
</tr>
<tr>
<td>4</td>
<td>1.0003</td>
<td>-216.04</td>
<td>198.42</td>
<td>-7679.40</td>
<td>6988.36</td>
<td>-91739.69</td>
<td>1.75x10^4</td>
<td>-1.44x10^4</td>
</tr>
<tr>
<td>5</td>
<td>1.0023</td>
<td>-215.69</td>
<td>197.48</td>
<td>-7606.26</td>
<td>6954.09</td>
<td>-91665.48</td>
<td>100</td>
<td>-1.21x10^4</td>
</tr>
<tr>
<td>6</td>
<td>1.0006</td>
<td>-216.01</td>
<td>199.76</td>
<td>-7612.32</td>
<td>7039.75</td>
<td>-91793.36</td>
<td>2.64x10^4</td>
<td>99.99</td>
</tr>
<tr>
<td>7</td>
<td>0.9999</td>
<td>-214.04</td>
<td>198.73</td>
<td>-7538.41</td>
<td>6999.29</td>
<td>-91739.50</td>
<td>100</td>
<td>99.99</td>
</tr>
</tbody>
</table>

![Figure 5-1](image1)

Real test
We used a piece of real image. The size of the image is 640x480 pixels, as illustrated in figure 5-2.

![Figure 5-2](image2)

Original image used for test

Table 5-4 calibration results of proposed method

<table>
<thead>
<tr>
<th>$X_s$ (cm)</th>
<th>$Y_s$ (cm)</th>
<th>$Z_s$ (cm)</th>
<th>$\phi$</th>
<th>$\omega$</th>
<th>$K$</th>
<th>$x_0$(pixel)</th>
<th>$z_0$(pixel)</th>
<th>$f$(pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-53.28</td>
<td>-62.01</td>
<td>21.75</td>
<td>1.3635</td>
<td>0.8825</td>
<td>-1.4452</td>
<td>-77.41</td>
<td>6.30</td>
<td>1628.24</td>
</tr>
</tbody>
</table>

Firstly, edges of objects (e.g. box) in the image are detected semi-automatically, we select corners, which are really in intersection point of two or more straight lines, as distinct points. The experimental results are shown in Table 5-4. Three rendered views of the reconstructed box is in figure 5-3.
figure 5-3 three rendered views of the reconstructed box

The advantages of this method are (1) The camera calibration was divided into two steps, this can obviously reduce the
rel entities among the orientation parameters. (2) During solving
the interior orientation parameters and rotation R, computational
process is linear, without any iteration and initiated values. (3)
The method makes full use of linear features and constraints (co-planar, parallel, vertical). It is robust and stable due to
its strong geometric and mathematical relations. The newest
trend to 3D reconstruction of building models consists of
exploiting information other than images to support the analysis,
for example fusing images with scanned, or digital maps. The
information about the ground plans of the building available in
the GIS map is used to define parametric building models.

References
1 G.Zhou. The intrinsic and extrinsic factors for CCD
calibration: distortion model and signal to noise ratio, Acta
2 H. Zielinski. Object reconstruction with digital line
photogrammetry. Doctoral thesis Department of Geodesy and
3 M.J. Magee and J.K. Aggarwal. Determining the position of a
robot using a single calibration object, Proc. IEEE Conf. on Robotics,
Atlanta, Georgia.
4 R.Y. Tsai. 1987 A versatile Camera Calibration Technique for
High Accuracy 3D Machine Vision Metrology Using Off-the-Shelf
TV Cameras and Lenses, IEEE J Robotics and Automation, vol. 3,
no.4 pp.323-331.
5 Tienan Chen, R. S. Shibasaki. 1998 DETERMINATION OF
CAMS ,S ORIENTATION PARAMETERS BASED ON LINE
FEATURES. International Archives of Photogrammetry and
6 T-Y Shih. The Duality and Critical Condition in the Formulation
and Decomposition of a Rotation Matrix, Photogrammetric and