Shirong Zheng

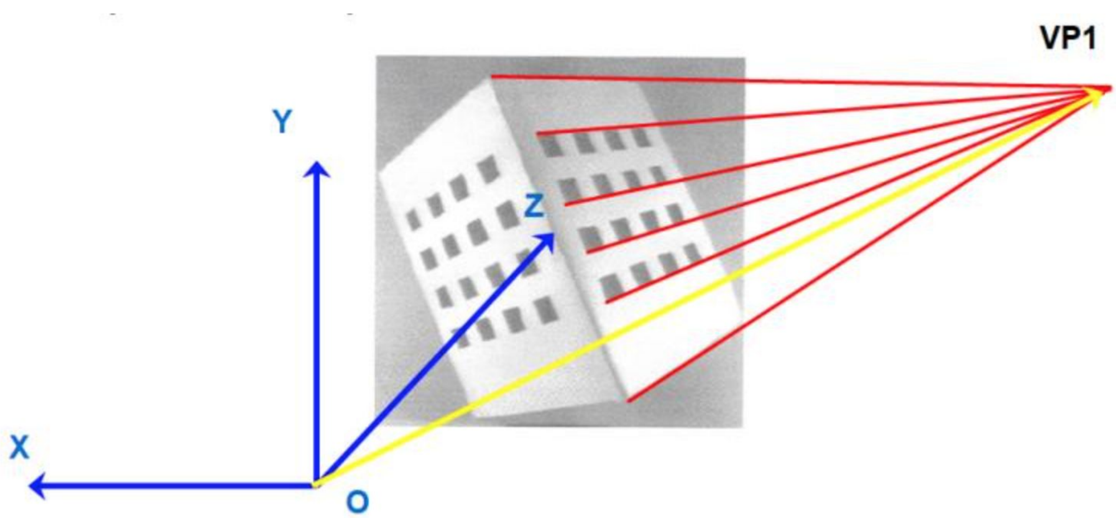
CSC47100-E

Prof. Zhu

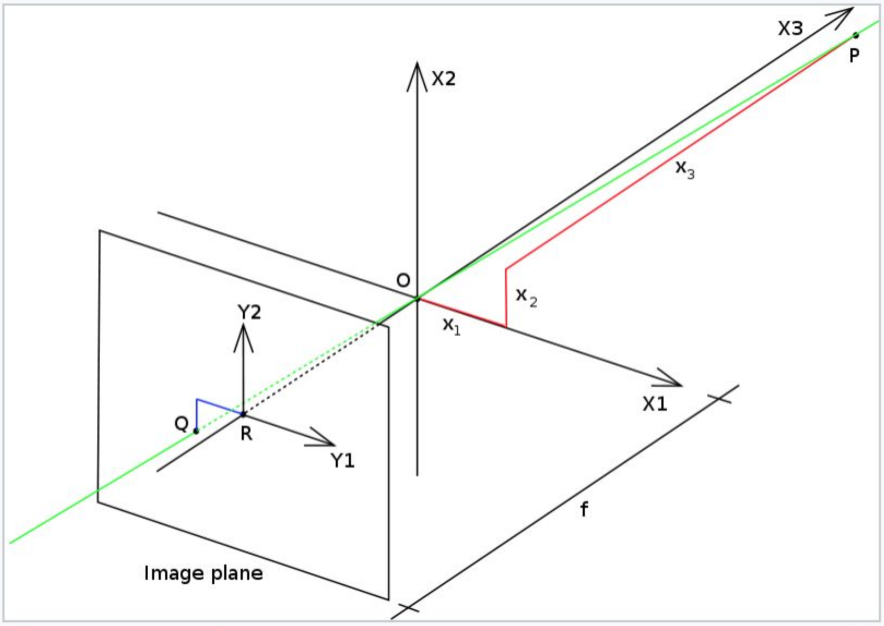
10/27/2017

Assignment 3

1. (Camera Models- 30 points) Prove that the vector from the viewpoint of a pinhole camera to the vanishing point (in the image plane) of a set of 3D parallel lines is parallel to the direction of the parallel lines. Please show steps of your proof.



The important property of camera model is Vector OV which is from the center of projection to the vanishing point is parallel to the parallel lines.

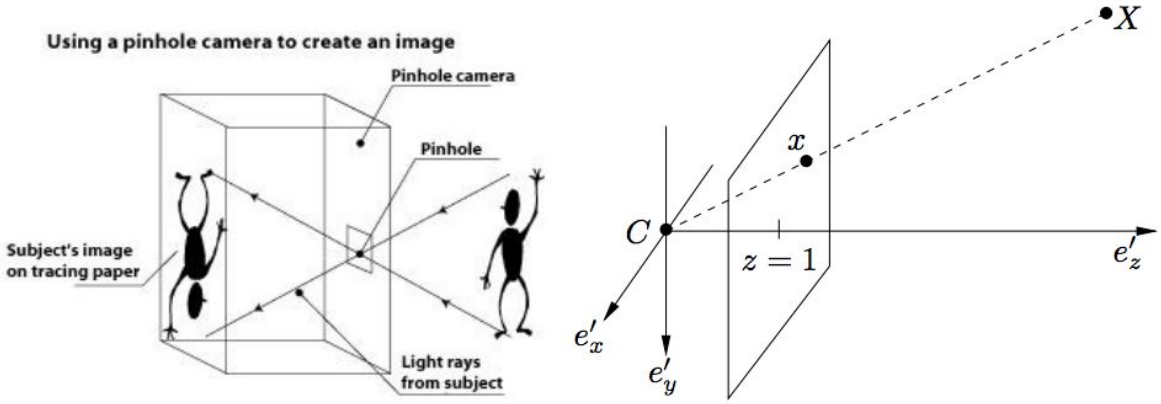


The geometry of a pinhole camera

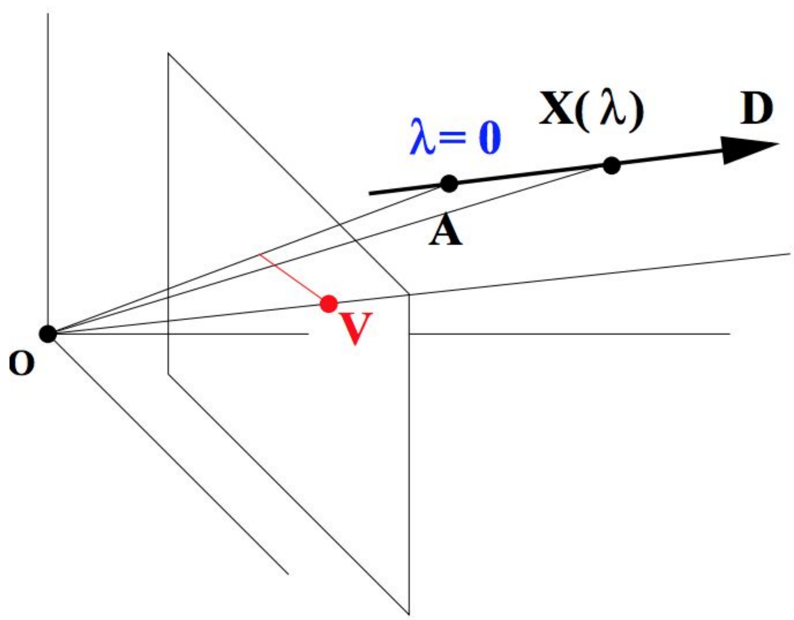
A 3D orthogonal coordinate system with its origin at O. This is also where the camera aperture is located. The three axes of the coordinate system are referred to as X1, X2, X3. Axis X3 is pointing in the viewing direction of the camera and is referred to as the optical axis, principal axis, or principal ray. The 3D plane which intersects with axes X1 and X2 is the front side of the camera, or principal plane.

A point R at the intersection of the optical axis and the image plane. This point is referred to as the principal point or image center. The projection of point P onto the image plane, denoted Q. This point is given by the intersection of the projection line (green) and the image plane.

There is a parallel line in the space, these parallel lines intersect at an infinity point, parallel to the image of the camera like the intersection of the infinity point of the image, that is, the disappearance of the point. The infinity and the camera are connected to the heart, will get a (red) with the family parallel straight line, and this line because of the light through the heart, it is like a solitary point, that is vanishing point.



Left is pinhole camera image, the right is the mathematical image



A line of 3D points isrepresented as $\mathbf{X}(\lambda) = \mathbf{A} + \lambda\mathbf{D}$

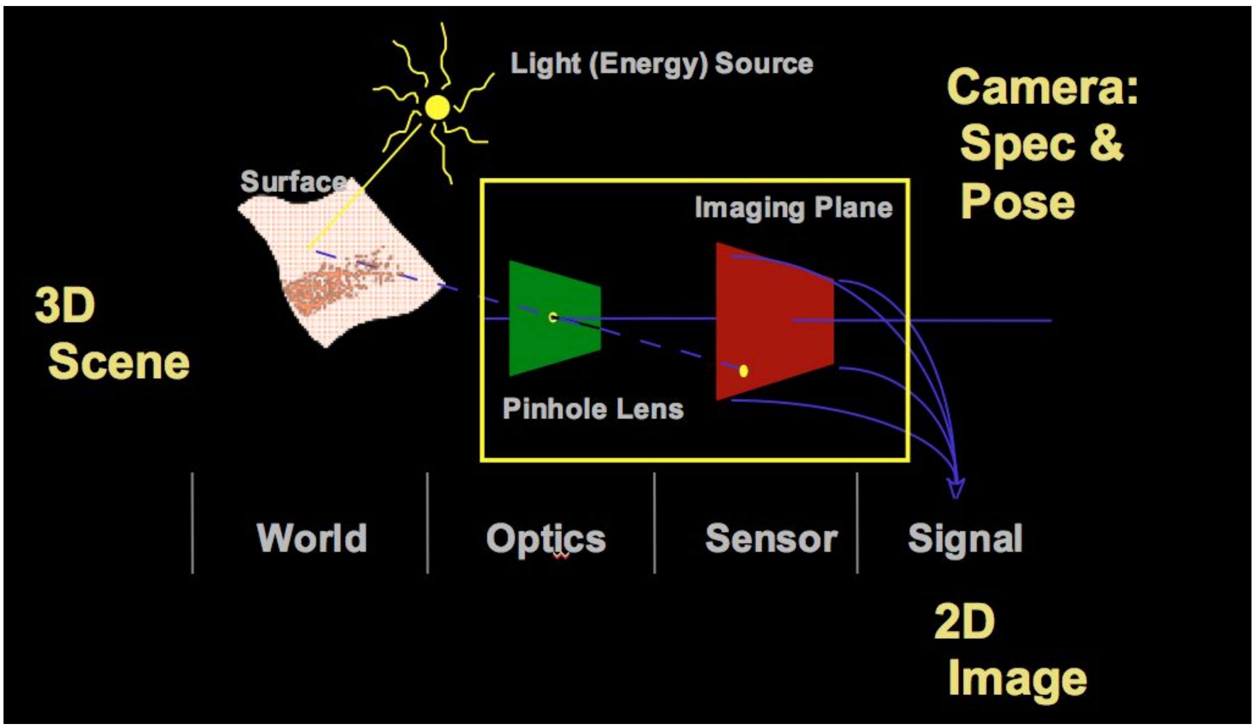
Using x=f(x/z), the vanishing point of its image is

$$\mathbf{v} = \lim_{\lambda \to \pm\infty} \mathbf{x}(\lambda) = f \frac{\mathbf{A} + \lambda \mathbf{D}}{A_Z + \lambda D_Z}$$

$$= f \frac{\mathbf{D}}{D_Z} = f \begin{pmatrix} D_X / D_Z \\ D_Y / D_Z \\ 1 \end{pmatrix}$$

v depends only on the direction D, not on A.
All parallel lines meet at the same vanishing point.

In other way we could say and use example from lecture



The blue line is the interpretation plane
If inter x pinhole = n1, imaging plane x inter=n2, if n1=n2 then it's parallel, if not then n1 dot product n2 =0, then it is parallel.

2. (Camera Models- 30 points) Show that relation between any image point $(xim, yim)^T$ of a plane (in the form of $(x1,x2,x3)^T$ in projective space) and its corresponding point $(Xw, Yw, Zw)^T$ on the plane in 3D space can be represented by a 3x3 matrix. You should start from the general form of the camera model $(x1,x2,x3)^T = M_{int}M_{ext}(Xw, Yw, Zw, 1)^T$, where the image center (ox, oy), the focal length f, the scaling factors(sx and sy), the rotation matrix R and the translation vector T are all unknown. Note that in the course slides and the lecture notes, I used a simplified model of the perspective project by assuming ox and oy are known and sx = sy =1, and only discussed the special cases of planes.. So you cannot directly copy those equations I used. Instead you should use the general form of the projective matrix, and the general form of a plane $n_x X_w + n_y Y_w + n_z Z_w = d$.

Shirong Zheng

#2

The relation between image plane to camera: $(x, y) = (\frac{fX}{Z}, \frac{fY}{Z})$

The relation between image plane and image frame:

$(x, y) = (-(x_{im} - o_x) s_x, -(y_{im} - o_y) s_y)$

∴ we could get the below equation

$(\frac{fX}{Z}, \frac{fY}{Z}) = (-(x_{im} - o_x) s_x, -(y_{im} - o_y) s_y)$

$$\begin{pmatrix} x_{im} \\ y_{im} \end{pmatrix} = \begin{pmatrix} (-\frac{fX}{s_x Z} + o_x) \\ (-\frac{fY}{s_y Z} + o_y) \end{pmatrix} = \begin{pmatrix} \left(\frac{-f_x X + o_x Z}{Z}\right) \\ \left(\frac{-f_y Y + o_y Z}{Z}\right) \end{pmatrix} = \begin{pmatrix} \left(\frac{-f_x X + 0Y + o_x Z}{0X + 0Y + Z}\right) \\ \left(\frac{0X - f_y Y + o_y Z}{0X + 0Y + Z}\right) \end{pmatrix}$$

∴ $\frac{f}{s_x} = f_x$, $\frac{f}{s_y} = f_y$

Camera Model Matrix (Image Frame)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -f_x & 0Y & o_x \\ 0X & -f_y & o_y \\ 0X & 0Y & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

World Coordinate System To Camera Coordinate System

$$P = RP_w + T \rightarrow \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}$$

In Camera Model, we convert to 3 x 4 matrix

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{33} & r_{33} & T_z \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

Combine the image frame ~~matrix~~ to Camera matrix and. the camera to world matrix.

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} -f_x & 0 & O_x \\ 0 & -f_y & O_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -f_x r_{11} + O_x r_{31} & -f_x r_{12} + O_x r_{32} & -f_x r_{13} + O_x r_{33} & -f_x T_x + O_x T_z \\ -f_y r_{21} + O_y r_{31} & -f_y r_{22} + O_y r_{32} & -f_y r_{23} + O_y r_{33} & -f_y T_y + O_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -f_x r_{11} X_w + O_x r_{31} X_w - f_x r_{12} Y_w + O_x r_{32} Y_w - f_x r_{13} Z_w + O_x r_{33} Z_w - f_x T_x + O_x T_z \\ -f_y r_{21} X_w + O_y r_{31} X_w - f_y r_{22} Y_w + O_y r_{32} Y_w - f_y r_{23} Z_w + O_y r_{33} Z_w - f_y T_y + O_y T_z \\ r_{31} X_w + r_{32} Y_w + r_{33} Z_w + T_z \end{pmatrix}$$

$$= \begin{pmatrix} (-f_x r_{11} + O_x r_{31}) X_w + (-f_x r_{12} + O_x r_{32}) Y_w + (-f_x r_{13} + O_x r_{33}) Z_w - f_x T_x + O_x T_z \\ (-f_y r_{21} + O_y r_{31}) X_w + (-f_y r_{22} + O_y r_{32}) Y_w + (-f_y r_{23} + O_y r_{33}) Z_w - f_y T_y + O_y T_z \\ r_{31} X_w + r_{32} Y_w + r_{33} Z_w + T_z \end{pmatrix}$$

$$\begin{pmatrix} X_{im} \\ Y_{im} \end{pmatrix}$$

$$= \begin{pmatrix} \dfrac{(-f_x r_{11} + O_x r_{31}) X_w + (-f_x r_{12} + O_x r_{32}) Y_w + (-f_x r_{13} + O_x r_{33}) Z_w - f_x T_x + O_x T_z}{r_{31} X_w + r_{32} Y_w + r_{33} Z_w + T_z} \\ \dfrac{(-f_y r_{21} + O_y r_{31}) X_w + (-f_y r_{22} + O_y r_{32}) Y_w + (-f_y r_{23} + O_y r_{33}) Z_w - f_y T_y + O_y T_z}{r_{31} X_w + r_{32} Y_w + r_{33} Z_w + T_z} \end{pmatrix}$$

By using a plane equation we add one more constraint for all the points on the plane:

$$n_x X_w + n_y Y_w + n_z Z_w = d$$

That could be written as

$$\mathbf{n}^{\mathbf{T}} \mathbf{P}_{\mathbf{w}} = d$$

Where n is the parameters of the plane, and Pw is a 3D point on the plane. Usually, Zw can be written as a function of Xw and Yw. For example, as a very special case, for the ground plane where a mobile robot moves, the plane equation is as simple as

$$Zw=0$$

Therefore we have Pw =(Xw,Yw,0,1)^T . The 3D point (Xw, Yw,Zw) becomes a 2D point(Xw, Yw), and we have the following model for a plane projection:

Assume this is equation Q1

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{bmatrix} -fr_{11} & -fr_{12} & -fT_x \\ -fr_{21} & -fr_{22} & -fT_y \\ r_{31} & r_{32} & T_z \end{bmatrix} \begin{pmatrix} X_w \\ Y_w \\ 1 \end{pmatrix}$$

We can assume nz = 1, therefore the plane equation can be written as:

$$Z_w = d - n_x X_w - n_y Y_w$$

We can easily derive the following 3x3 transformation matrix for the more general case by plugging the above equation into equation Q1, with the M matrix in Eq.

$$\mathbf{M} = \begin{bmatrix} -fr_{11} & -fr_{12} & -fr_{13} & -fT_x \\ -fr_{21} & -fr_{22} & -fr_{23} & -fT_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

Assume this is equation Q2

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{bmatrix} -f(r_{11} - n_x r_{13}) & -f(r_{12} - n_y r_{13}) & -f(dr_{13} + T_x) \\ -f(r_{21} - n_x r_{23}) & -f(r_{22} - n_y r_{23}) & -f(dr_{23} + T_y) \\ (r_{31} - n_x r_{33}) & (r_{32} - n_y r_{33}) & (dr_{33} + T_z) \end{bmatrix} \begin{pmatrix} X_w \\ Y_w \\ 1 \end{pmatrix}$$

After that, for any point on the plane, we can calculate its 3D coordinates in the world coordinate system. In that sense, we can obtain depths of points of a 3D plane from a single image.

Then find out Mext which is a 3x4 matrix that only include extrinsic parameters

$$\mathbf{M}_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1^T & T_x \\ \mathbf{R}_2^T & T_y \\ \mathbf{R}_3^T & T_z \end{bmatrix}$$

Mint is a 3x3 matrix that includes only the intrinsic parameters

$$\mathbf{M}_{\text{int}} = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

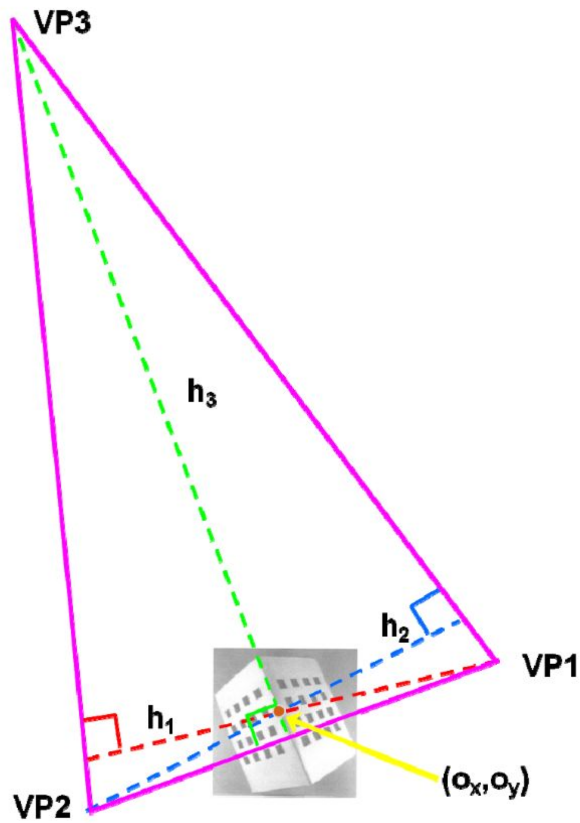
Combine Mext and Mint then substitute into equation Q1, we finally have a linear representation of the perspective projection transformation ast

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{M}_{\text{int}}\mathbf{M}_{\text{ext}} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

Each side divide by X3, that we define (x1,x2,x3)T as below.

$$\begin{pmatrix} x_{im} \\ y_{im} \end{pmatrix} = \begin{pmatrix} x_1 / x_3 \\ x_2 / x_3 \end{pmatrix}$$

3. (Calibration- 40 points) Prove the Orthocenter Theorem by geometric arguments: Let T be the triangle on the image plane defined by the three vanishing points of three mutually orthogonal sets of parallel lines in space. Then the image center is the orthocenter of the triangle T (i.e., the common intersection of the three altitudes.
(1) Basic proof: use the result of Question 1, assuming the aspect ratio of the camera is 1. (20 points)
(2) If you do not know the focal length of the camera, can you still find the image center (together with the focal length) using the Orthocenter Theorem? Show why or why not. (10 points)
(3) If you do not know the aspect ratio and the focal length of the camera, can you still find the image center using the Orthocenter Theorem? Show why or why not. (10 points)



Prove:
The center of projection of the camera in 3D space is O. Three mutually orthogonal sets of parallel lines: L1, L2, and L3. Assume there is a formed triangle V1V2V3, then V1, V2, and V3 are the three vanishing points.

(1)Basic proof: We had prove that the vector OVi from the viewpoint or pinhole to a vanishing point Vi is parallel to the mutually orthogonal set of parallel lines Li in 3D space (i=1,2,3) in question 1.

∵ OVi ⊥ OVj but i≠j

∴ OV1 ⊥ V2V3, OV2 ⊥ V1V3, OV3 ⊥V1V2.

∵ Vihi is the altitude from Vi (i=1,2,3)

∴V1h1 ⊥ V2V3, V2h2 ⊥ V1V3, V3h3 ⊥ V1V2.

∵o is center of the image,

∴line Oo ⊥ the image plane. The point o is the projection of point O in the image plane, which lies on the all 3 altitudes Vihi, so o is the intersection point of V1h1, V2h2, and V3h3.

(2)There is nothing about any information of the camera parameter to find the vanishing points. We even don't use the focal length information, the only thing we need is the orthogonal relations of the altitude and aspect ratio of the image. That required to obtained a scaled image in the y direction if aspect ration was wrong. The vanishing points would still be correct, but the image center will not be at (ox, oy), since hi is not the altitude any more.

(3) The orthocenter theorem only works on images without lens distortions or with lens distortions removed(there are unknown of aspect ratio and the focal length).

If the camera ⊥ the plane (the set of parallel lines), then the images of these parallel lines would be still in parallel and would not have an intersection. In this three sets of parallel lines that are mutually orthogonal to each other, so there are three vanishing points generated from t. Therefore, their vanishing point will be at infinity. In this case we cannot form a triangle to estimate the image center.