Mid-Term

Honor Pledge: I pledge on my honor that I have not given or received any unauthorized assistance on this assignment/examination.

1. Question 3

(a)

For any given plane $\mathbf{a} \cdot \mathbf{v} = 1$, the normal is $\mathbf{a}$. Hence, its unit normal is $\frac{\mathbf{a}}{|\mathbf{a}|}$.

Recall the distance of a point p to the plane $\mathbf{a} \cdot \mathbf{v} - 1 = 0$ is computed as $\frac{|\mathbf{a} \cdot p - 1|}{|\mathbf{a}|}$. Hence, the distance of the plane from the origin (i.e., $p = (0, 0, 0)$) is $\frac{1}{\sqrt{\mathbf{a} \cdot \mathbf{a}}}$.

(b)

The line joining two points is corresponding to the intersection line of two plane in the dual space. We know the normal of a plane $\mathbf{a} \cdot \mathbf{v} = 1$ is simply $\mathbf{a}$. Hence, the normals for plane $\mathbf{a_1} \cdot \mathbf{v} = 1$ and $\mathbf{a_2} \cdot \mathbf{v} = 1$ is $\mathbf{a_1}$ and $\mathbf{a_2}$ respectively. The cross product of two normals would result in a direction which is perpendicular to both normals. This vector, $\mathbf{a_1} \times \mathbf{a_2}$ is exactly the direction of the intersecting line L of those two planes. Now, we only need a point on L to get the equation of it.
Let us assume $\mathbf{a_1} = (\alpha_1, \beta_1, \gamma_1)$ and $\mathbf{a_2} = (\alpha_2, \beta_2, \gamma_2)$. We have the following equations set:

$$\begin{aligned} \alpha_1 v_1 + \beta_1 v_2 + \gamma_1 v_3 &= 1 \\ \alpha_2 v_1 + \beta_2 v_2 + \gamma_2 v_3 &= 1. \end{aligned}$$

Let us set $v_1 = 0$, then we can compute $v_2 = \frac{\gamma_2\beta_1 - \gamma_1\beta_2}{\gamma_2 - \gamma_1}$ and $v_3 = \frac{\beta_2\gamma_1 - \beta_1\gamma_2}{\beta_2 - \beta_1}$. Hence, we found a point $(0, \frac{\gamma_2\beta_1 - \gamma_1\beta_2}{\gamma_2 - \gamma_1}, \frac{\beta_2\gamma_1 - \beta_1\gamma_2}{\beta_2 - \beta_1})$ on the intersecting line L. Therefore,

$$L : (0, \frac{\gamma_2\beta_1 - \gamma_1\beta_2}{\gamma_2 - \gamma_1}, \frac{\beta_2\gamma_1 - \beta_1\gamma_2}{\beta_2 - \beta_1}) + \lambda(\mathbf{a_1} \times \mathbf{a_2}).$$

The dual of the line formed by the intersection of the two planes $\mathbf{a_1} \cdot \mathbf{v} = 1$ and $\mathbf{a_2} \cdot \mathbf{v} = 1$ is simply the line joining two points $\mathbf{a_1}$ and $\mathbf{a_2}$, which has the following equation

$$L : \mathbf{a_1} + \lambda(\mathbf{a_1} \times \mathbf{a_2}).$$

(c)

After superimpose dual space on object space, we have the following two lines:

$$\begin{aligned} L_1 :&(0, \frac{\gamma_2\beta_1 - \gamma_1\beta_2}{\gamma_2 - \gamma_1}, \frac{\beta_2\gamma_1 - \beta_1\gamma_2}{\beta_2 - \beta_1}) + \lambda(\mathbf{a_1} \times \mathbf{a_2}) \\ L_2 :&\mathbf{a_1} + \lambda(\mathbf{a_1} \times \mathbf{a_2}) \end{aligned}$$

Their normal is hence $\mathbf{a_1} \times \mathbf{a_2}$ and $\mathbf{a_1} - \mathbf{a_2}$ respectively. We know $\mathbf{a_1} - \mathbf{a_2}$ is collinear (i.e., parallel) with $\mathbf{a_1}$ or $\mathbf{a_2}$. Thus, we could conclude that $(\mathbf{a_1} \times \mathbf{a_2}) \cdot (\mathbf{a_1} - \mathbf{a_2}) = 0$, i.e., L_1 and L_2 are perpendicular to each other.

2. Question 5

(a)

For any given ellipse in a plane, we could always form a cone together with the camera center. As illustrated in Figure **??**, camera center is shown at the top corner of the cone. In order to proper form the image, the image plane should **not** be parallel to any generating line nor the center axis of that cone, i.e., the cases of parabola or hyperbola in Figure **??** would not happen, we would be able to conclude that the image on the image plane of any ellipse is always an ellipse.

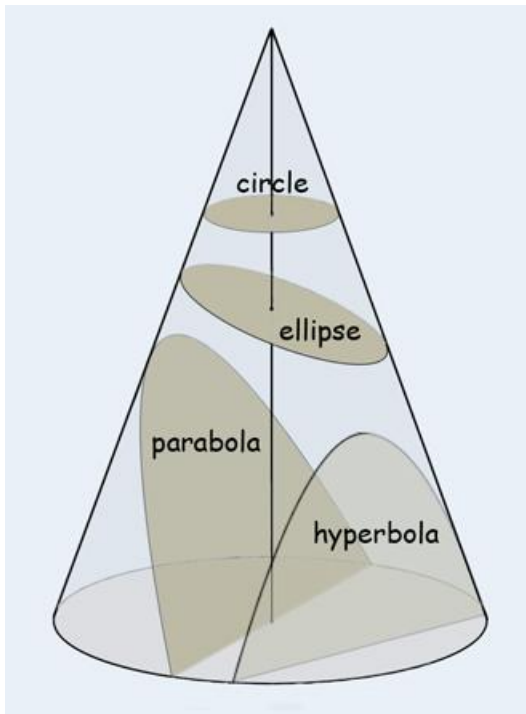

Figure 1: Conic Section

(b)

From Figure **??**, we could always replace the sphere by a circle C, i.e., an intersection plane, of radius r_c, and get the same image on the image plane. Given the values of sphere radius r and distance from the camera center d, we could compute the r_c as $r_c = \frac{r\sqrt{d^2 - r^2}}{d}$. The distance of the center of the circle C from the origin of the camera reference frame is given by $\delta = \sqrt{d^2 - 2r^2 + r^2(\frac{r}{d})^2}$. Now, let us consider the plane P which the circle C lies on. Let us define two unit mutually orthogonal vectors $\vec{a} = (a_1, a_2, a_3)^T$ and $\vec{b} = (b_1, b_2, b_3)^T$ in plane P. Simply, we define the vector $\vec{c} = (c_1, c_2, c_3)^T$ as the unit normal to plane P. Hence, if we have (X, Y, Z) with respect to the camera reference frame, its corresponding representation with respect to P's coordination system is

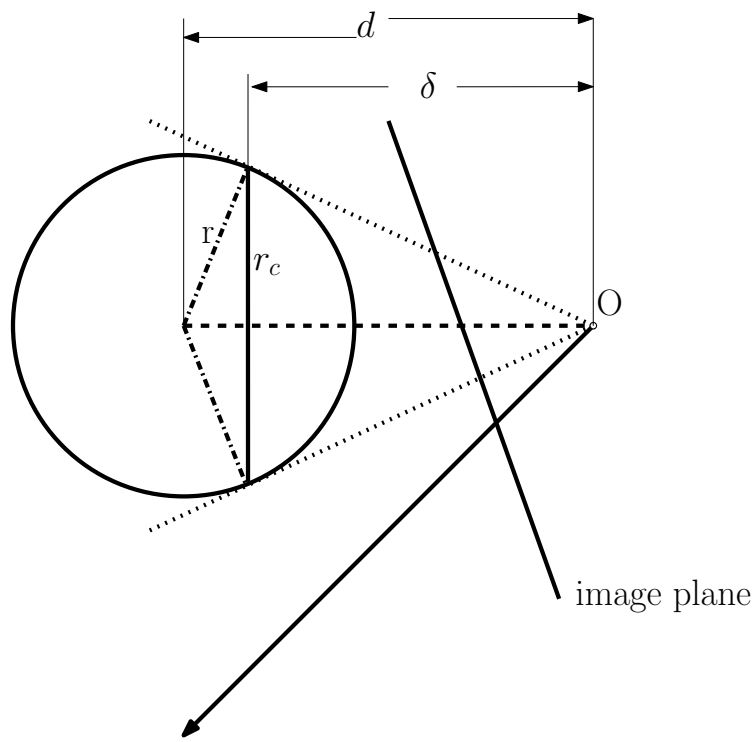

Figure 2: Sphere Perspective Projection

(X', Y', Z') is related as

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} \vec{a} \\ \vec{b} \\ \vec{c} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

A point (X, Y, Z) in the camera reference frame and its perspective projection(x, y) on the image plane is related by

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} x \\ y \end{pmatrix}$$

where f is the focal length of the camera. Hence,

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} \vec{a} \\ \vec{b} \\ \vec{c} \end{pmatrix} \begin{pmatrix} x \\ y \\ f \end{pmatrix}$$

From this equation, we get the relationship of the projected coordinates and the coordinates system of P is

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \frac{\delta}{c_1 x + c_2 y + c_3 f} \begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix} \begin{pmatrix} x \\ y \\ f \end{pmatrix}$$

To be more explicit,

$$(c_1 x + c_2 y + c_3 f) X' = \delta (a_1 x + a_2 y + a_3 f)$$
$$(c_1 x + c_2 y + c_3 f) Y' = \delta (b_1 x + b_2 y + b_3 f)$$

Since we know a circle on plane P can be represented in the plane coordinate frame as

$$X'^2 + Y'^2 = r_c^2$$

The circle is centered at the origin of plane P with radius r_c. By substitutin, we will have

$$\delta^2(a_1x + a_2y + a_3f)^2 + \delta^2(b_1x + b_2y + b_3f)^2 = (c_1x + c_2y + c_3f)^2 r_c^2$$

Rearrange all the terms, we will have

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

where

$$\begin{aligned}
A &= \delta^2 - c_1^2(\delta^2 + r_c^2) \\
B &= -c_1c_2(\delta^2 + r_c^2) \\
C &= \delta^2 - c_2^2(\delta^2 + r_c^2) \\
D &= -c_1c_3(\delta^2 + r_c^2)f \\
E &= -c_2c_3(\delta^2 + r_c^2)f \\
F &= (\delta^2 - c_3^2(\delta^2 + r_c^2))f^2
\end{aligned}$$

which follows the expression of an ellipse.
The slope of the major axis of the ellipse is given by

$$\begin{aligned}
m &= \frac{2B}{A - C - \sqrt{(A-C)^2 + 4B^2}} \\
&= \frac{-2c_1c_2(\delta^2 + r_c^2)}{-c_1^2(\delta^2 + r_c^2) + c_2^2(\delta^2 + r_c^2) - \sqrt{(-c_1^2(\delta^2 + r_c^2) + c_2^2(\delta^2 + r_c^2))^2 + 4(c_1c_2(\delta^2 + r_c^2))^2}} \\
&= \frac{-2c_1c_2}{-c_1^2 + c_2^2 - \sqrt{(-c_1^2 + c_2^2)^2 + 4(c_1c_2)^2}} \\
&= \frac{c_2}{c_1}
\end{aligned}$$

The center point of the ellipse is

$$\begin{aligned}
x_c &= -\frac{c_1c_3d^2f}{c_3^2d^2 - r^2} \\
y_c &= -\frac{c_2c_3d^2f}{c_3^2d^2 - r^2}
\end{aligned}$$

Hence, the equation of the major axis of the ellipse in the image plane is

$$\begin{aligned}
y &= m(x - x_c) + y_c \\
&= \frac{c_2}{c_1}(x + \frac{c_1c_3d^2f}{c_3^2d^2 - r^2}) - \frac{c_2c_3d^2f}{c_3^2d^2 - r^2}
\end{aligned}$$

Rearrange the terms, we will have

$$(c_3^2d^2 - r^2)y = \frac{c_2}{c_1}((c_3^2d^2 - r^2)x + c_1c_3d^2f) - c_2c_3d^2f$$

Clearly, the origin of the image $(0, 0)$ would satisfy the above equation, hence. Therefore, we have proved that the image of a sphere is an ellipse and its major axis would pass through the origin of the image.
Hence, given one particular sphere, we could always fit an ellipse on the image plane. After getting all the parameters of that ellipse, we could always find such a line without any knowledge on its normal nor its distance from camera center. Hence, if we could locate the optical center of a camera from their projected ellipses. First compute the major axis of their images, then find the intersection point. If we are given three or more spheres, we could optimize the result via least square minimization techniques.

3. Question 6

Let us assume the center point of polyhedron image and the camera center forms the projection axis. In this case, we take the unit vector $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, $\beta = (\beta_1, \beta_2, \beta_3)$ and $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, as illustrated in Figure **??**.

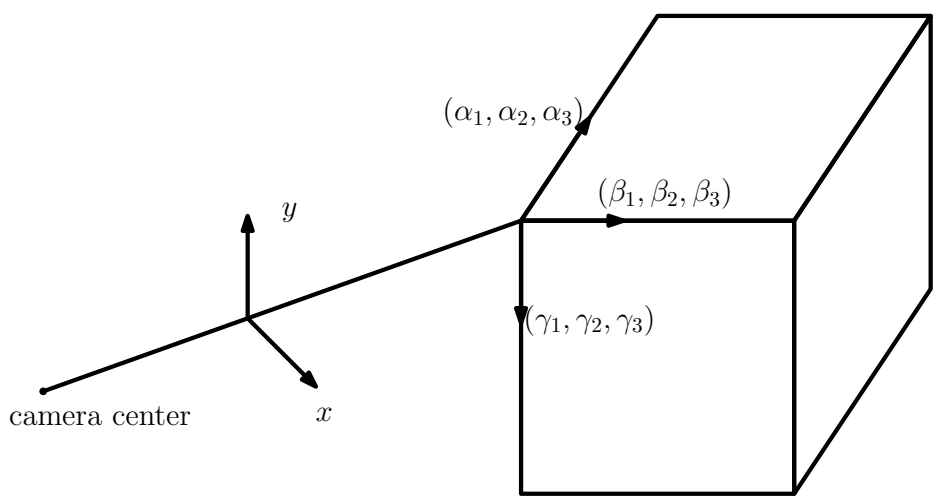

Figure 3: Polyhedron Projection

To get the orientation of the polyhedron, we need to estimate these three vectors, i.e., 9 unknowns. Their natural properties provide us three equations.

$$|\alpha| = 1$$
$$|\beta| = 1$$
$$|\gamma| = 1$$

We know in 3D space, these three vectors are orthogonal to each other, i.e., their dot product is 0. This provides us another three equations.

$$\alpha \cdot \beta = 0$$
$$\beta \cdot \gamma = 0$$
$$\gamma \cdot \alpha = 0$$

We have their 2D projection as $(\frac{\alpha_1}{\alpha_3}f, \frac{\alpha_2}{\alpha_3}f)$, $(\frac{\beta_1}{\beta_3}f, \frac{\beta_2}{\beta_3}f)$ and $(\frac{\gamma_1}{\gamma_3}f, \frac{\gamma_2}{\gamma_3}f)$. Since we know their angles in the image, these provides us another three equations. Let us assume the image was formed at focal length $f = 1$.

$$\begin{aligned}
\frac{\alpha_1}{\alpha_3}\frac{\beta_1}{\beta_3} + \frac{\alpha_2}{\alpha_3}\frac{\beta_2}{\beta_3} &= \cos\theta_1 \\
\frac{\beta_1}{\beta_3}\frac{\gamma_1}{\gamma_3} + \frac{\beta_2}{\beta_3}\frac{\gamma_2}{\gamma_3} &= \cos\theta_2 \\
\frac{\alpha_1}{\alpha_3}\frac{\gamma_1}{\gamma_3} + \frac{\alpha_2}{\alpha_3}\frac{\gamma_2}{\gamma_3} &= \cos\theta_3
\end{aligned}$$

Hence, we could be able to solve three vectors and get the orientation of the polyhedron. Take note that, we assume the focal length $f = 1$ just to get the orientation. The angle remains same as we form the image at different focal length or move the polyhedron along the axis closer or further — it is the scale that changes. However, if we want to know the exact distance of the polyhedron to the camera, we need a lot more information.

4. Question 8

(a)

Let us assume the camera is located at the origin of the world coordinate syatem, and there are no skew, translation nor distortion, we will have the following equations under homegraphic coodinates.

$$\begin{aligned}
Z &= Z_0 + pX + qY \\
1 &= \frac{Z_0}{Z} + p\frac{X}{Z} + q\frac{Y}{Z} \\
\frac{Z_0}{Z} &= 1 - px - qy,
\end{aligned}$$

where $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$, and (X, Y, Z) are the 3D space coordinates and (x, y) are the 2D image coordinates of the respective point.

(b)

The equations for motion field are

$$\begin{aligned}
u &= \frac{Wx - Uf}{Z} + \frac{\alpha xy}{f} - \beta(f + \frac{x^2}{f}) + \gamma y \\
v &= \frac{Wy - Vf}{Z} + \alpha(f + \frac{x^2}{f}) - \frac{\beta xy}{f} - \gamma x
\end{aligned}$$

Since $\frac{1}{Z} = \frac{1 - px - qy}{Z_0}$ and the focal length $f = 1$, we have

$$\begin{aligned}
u &= \frac{1}{Z_0}(Wx - U)(1 - px - qy) + \alpha xy - \beta(1 + x^2) + \gamma y \\
v &= \frac{1}{Z_0}(Wy - V)(1 - px - qy) + \alpha(1 + x^2) - \beta xy - \gamma x
\end{aligned}$$

(c)

If we are given two sets of different rigid motions with respect to different planar surfaces, $(\mathbf{t_1}, \omega_\mathbf{1}, Z_{0_1}, p_1, q_1)$ and $(\mathbf{t_2}, \omega_\mathbf{2}, Z_{0_2}, p_2, q_2)$, Hence,

$$u_1 = u_2$$
$$v_1 = v_2$$

and we have

$$\frac{1}{Z_{0_1}}(W_1x - U_1)(1 - p_1x - q_1y) + \alpha_1xy - \beta_1(1 + x^2) + \gamma_1y$$
$$= \frac{1}{Z_{0_2}}(W_2x - U_2)(1 - p_2x - q_2y) + \alpha_2xy - \beta_2(1 + x^2) + \gamma_2y$$
$$\frac{1}{Z_{0_1}}(W_1x - U_1 - p_1W_1x^2 + p_1U_1x - q_1W_1xy + q_1U_1y) + \alpha_1xy - \beta_1 - \beta_1x^2 + \gamma_1y$$
$$= \frac{1}{Z_{0_2}}(W_2x - U_2 - p_2W_2x^2 + p_2U_2x - q_2W_2xy + q_2U_2y) + \alpha_2xy - \beta_2 + \beta_2x^2 + \gamma_2y$$

Similarly for v, we have

$$\frac{1}{Z_{0_1}}(W_1y - V_1)(1 - p_1x - q_1y) + \alpha_1(1 + y^2) - \beta_1xy - \gamma_1x$$
$$= \frac{1}{Z_{0_2}}(W_2y - V_2)(1 - p_2x - q_2y) + \alpha_2(1 + y^2) - \beta_2xy - \gamma_2x$$
$$\frac{1}{Z_{0_1}}(W_1y - V_1 - p_1W_1xy + p_1V_1x - q_1W_1y^2 + q_1V_1y) + \alpha_1 + \alpha_1y^2 - \beta_1xy - \gamma_1x$$
$$= \frac{1}{Z_{0_2}}(W_2y - V_2 - p_2W_2xy + p_2V_2x - q_2W_2y^2 + q_2V_2y) + \alpha_2 + \alpha_2y^2 - \beta_2xy - \gamma_2x$$

By equating all the corresponding coefficients, we have

$$\frac{U_1}{Z_{0_1}} + \beta_1 = \frac{U_2}{Z_{0_2}} + \beta_2$$
$$\frac{W_1 + p_1U_1}{Z_{0_1}} = \frac{W_2 + p_2U_2}{Z_{0_2}}$$
$$\frac{q_1U_1}{Z_{0_1}} + \gamma_1 = \frac{q_2U_2}{Z_{0_2}} + \gamma_2$$
$$\frac{q_1W_1}{Z_{0_1}} - \alpha_1 = \frac{q_2W_2}{Z_{0_2}} - \alpha_2$$
$$\frac{p_1W_1}{Z_{0_1}} + \beta_1 = \frac{p_2W_2}{Z_{0_2}} + \beta_2$$
$$\frac{V_1}{Z_{0_1}} - \alpha_1 = \frac{V_2}{Z_{0_2}} - \alpha_2$$
$$\frac{p_1V_1}{Z_{0_1}} - \gamma_1 = \frac{p_2V_2}{Z_{0_2}} - \gamma_2$$
$$\frac{W_1 + q_1V_1}{Z_{0_1}} = \frac{W_2 + q_2V_2}{Z_{0_2}}$$
$$\frac{p_1W_1}{Z_{0_1}} + \beta_1 = \frac{p_2W_2}{Z_{0_2}} + \beta_2$$
$$\frac{q_1W_1}{Z_{0_1}} - \alpha_1 = \frac{q_2W_2}{Z_{0_2}} - \alpha_2$$

5. Question 10

2	3	4	5	6
3	4	5	6	8
4	5	6	8	5
5	7	8	9	3
9	10	9	4	3

(a)

5.1.1 A 3×3 Gaussian Filter

-	-	-	-	-
-	4	5	6	-
-	5	6	8	-
-	7	8	7	-
-	-	-	-	-

5.1.2 A 3×3 Box Filter

-	-	-	-	-
-	4	5	6	-
-	5	6	6	-
-	7	7	6	-
-	-	-	-	-

(b)

After applying a 3×3 Sobel horizontal edge detector, we have

-	-	-	-	-
-	1.0	1.125	0.875	-
-	1.125	1.25	-0.5	-
-	1.0	0.125	-2.125	-
-	-	-	-	-

After applying a 3×3 Sobel vertical edge detector, we have

-	-	-	-	-
-	-1.0	-1.125	-0.875	-
-	-1.375	-1.5	-0.5	-
-	-2.25	-0.875	0.875	-
-	-	-	-	-

Therefore, the gradient along the vertical direction is 1.25, while the gradient along the horizontal direction is 1.5.[1] As a result, the tangent of the edge with respect to the horizontal axis would be $\frac{1.25}{1.5}$ — the angle formed by the edge and the horizontal axis would be $\arctan \frac{1.25}{1.5} \approx 50.1944°$.

[1] We take the absolute value for the gradient.

(c)

-	-	-	-	-
-	4.0	5.0	5.5	-
-	5.0	6.0	6.0	-
-	7.0	8.0	6.0	-
-	-	-	-	-

The median filter works best for salt and pepper noise, as the noise lies in two extremas (i.e., either very small or very large) of the intensity distribution. Hence, it would not greatly affect the median.

(d)

First, Gaussian filter defines a kernel which gives less weight to pixels further form the center of the window, i.e., it weights nearby points more than distant ones. Second, Gaussian filter is rotationally symmetric, i.e., it behaves as a circularly symmetric fuzzy bulb. This makes Gaussian filter a good smoothing filter.
Gaussian filter could be implemented in a fast manner. Due to the additive property, we could break a Guassian filter (as well as a Box filter) into two steps — first convolve each row with a 1-D filter, and the convolve each column with a 1-D filter. Equivalently, we could implement it in one operation — convolve in rows, and transpose the underlying image if we want to convolve columns.

(e)

If we apply a Gaussian filter with a kernel size larger than the width of the line, we will end up with the case that the detected edges drift aparts from each other. This is because edge detector would always look at the point with largest first derivative. A Gaussian filter with a large kernel size would smooth the underlying line and make two edges drift apart. This would certainly introduce error during edge detection. This is also the reason behind many optical illusion.

(f)

The Box filter defines the intensity of a point as the pixel intensity average of a certain area around it. As a result, under the assumption that **noise remains minority in any given specified area** (which is usually granted), it would attenuate noise as it smooth "out" the noise.