

Appendices



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Appendix A

Camera Models and Projective Camera Anatomy

A.1 CCD Cameras

The pinhole camera model just derived assumes that the image coordinates are Euclidean coordinates having equal scales in both axial directions. In the case of Charge coupled device (CCD) cameras, there is the additional possibility of having non-square pixels. If image coordinates are measured in pixels, then this has the extra effect of introducing unequal scale factors in each direction. In particular if the number of pixels per unit distance in image coordinates are m_x and m_y in the x and y directions, then the transformation from world coordinates to pixel coordinates is obtained by multiplying 3.4 on the left by an extra factor $\text{diag}(m_x, m_y, 1)$. Thus, the general form of the calibration matrix of a CCD camera is



University of Moratuwa, Sri Lanka
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

$$K = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \quad (\text{A.1})$$

where $\alpha_x = fm_x$ and $\alpha_y = fm_y$ represent the focal length of the camera in terms of pixel dimensions in the x and y direction respectively. Similarly, $\tilde{\mathbf{x}}_0 = (x_0, y_0)$ is the principal point in terms of pixel dimensions, with coordinates $x_0 = m_x p_x$ and $y_0 = m_y p_y$. A CCD camera thus has 10 degrees of freedom.

A.2 Finite Projective Camera

For added generality, we can consider a calibration matrix of the form

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \quad (\text{A.2})$$

The added parameter s is referred to as the skew parameter. The skew parameter will be zero for most normal cameras. A camera

$$P = KR[I \mid -\vec{C}] \quad (\text{A.3})$$

for which the calibration matrix K is of the form A.2 will be called a finite projective camera. A finite projective camera has 11 degrees of freedom. This is the same number of degrees of freedom as a 3×4 matrix, defined up to an arbitrary scale.

Note that the left hand 3×3 sub-matrix of P , equal to KR , is non-singular. Conversely, any 3×4 matrix P for which the left hand 3×3 sub-matrix is non-singular is the camera matrix of some finite projective camera, because P can be decomposed as $P = KR[I \mid -\vec{C}]$. Indeed, letting M be the left 3×3 sub-matrix of P , one decomposes M as a product $M = KR$ where K is upper-triangular of the form A.2 and R is a rotation matrix. This decomposition is essentially the RQ matrix decomposition, described in Appendix C. The matrix P can therefore be written $P = M[I \mid M^{-1}\mathbf{p}_4] = KR[I \mid -\vec{C}]$ where \mathbf{p}_4 is the last column of P . In short

- The set of camera matrices of finite projective cameras is identical with the set of homogeneous 3×4 matrices for which the left hand 3×3 sub-matrix is non-singular.



University of Moratuwa, Sri Lanka.

Electronic Theses & Dissertations

www.lib.mrt.ac.lk

The Table A.1 provides a summary of the properties of projective camera P .

Table A.1: Summary of the properties of projective camera P. The matrix is represented by the block form $P = [M \mid \vec{p}_4]$.

| |
|--|
| <p>Camera center: The camera center is the 1-dimensional right null-space \vec{C} of P, i.e. $P\vec{C} = 0$.</p> <p>Finite camera:(M is not singular) $\vec{C} = \begin{pmatrix} -M^{-1}\vec{p}_4 \\ 1 \end{pmatrix}$</p> <p>Camera at infinity:(M is singular) $\vec{C} = \begin{pmatrix} \vec{d} \\ 0 \end{pmatrix}$ where \vec{d} is the null 3-vector of M, i.e. $M\vec{d} = 0$.</p> <p>Column points: For $i = 1, \dots, 3$, the column vectors \vec{p}_i are vanishing points in the image corresponding to the X, Y and Z axes respectively. Column \vec{p}_4 is the image of the coordinate origin.</p> <p>Principal plane: The principal plane of the camera is \mathbf{P}^3, the last row of P.</p> <p>Axis planes: The planes \mathbf{P}^1 and \mathbf{P}^2 (the first and second rows of P) represent planes in space through the camera center, corresponding to points that map to the image lines $x = 0$ and $y = 0$ respectively.</p> <p>Principal point: The image point $x_0 = M\vec{m}^3$ is the principal point of the camera, where \vec{m}^{3T} is the third row of M.</p> <p>Principal ray: The principal ray (axis) of the camera is the ray passing through the camera center \vec{C} with direction vector \vec{m}^{3T}. The principal axis vector $\vec{v} = \det(M)\vec{m}^3$ is directed towards the front of the camera.</p> |
|--|



University of Moratuwa, Sri Lanka
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

A general projective camera P maps world points \vec{X} to image points \vec{x} according to $\vec{x} = P\vec{X}$. Building on this mapping we will now dissect the camera model to reveal how geometric entities, such as the camera center, are encoded. Some of the properties that we consider will apply only to finite projective cameras and their specializations, whilst others will apply to general cameras. The distinction will be evident from the context. The derived properties of the camera are summarized in Table A.1.

A general projective camera may be decomposed into blocks according to $P = [M \mid \vec{p}_4]$, where M is a 3×3 matrix. It will be seen that if M is non-singular, then this is a finite camera, otherwise it is not.

A.3 Camera center

The matrix P has a 1-dimensional right null-space because its rank is 3, whereas it has 4 columns. Suppose the null-space is generated by the 4-vector \vec{C} , that is $P\vec{C} = 0$. It will now be shown that \vec{C} is the camera center, represented as a homogeneous 4-vector.

Consider the line containing \vec{C} and any other point \vec{A} in 3-space. Points on this line may be represented by the join

$$\vec{X}(\lambda) = \lambda\vec{A} + (1 - \lambda)\vec{C}.$$

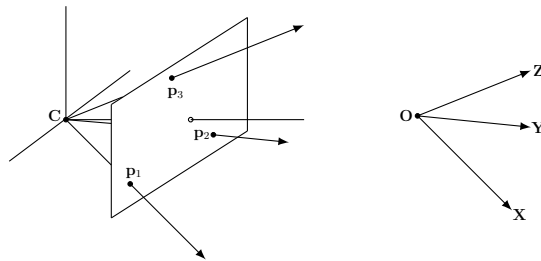


Figure A.1: The three image points defined by the columns \mathbf{p}_i , $i = 1, \dots, 3$, of the projection matrix are the vanishing points of the directions of the world axes.

Under the mapping $\vec{x} = P\vec{X}$ points on this line are projected to

$$\vec{x} = P\vec{X}(\lambda) = \lambda P\vec{A} + (1 - \lambda)P\vec{C} = \lambda P\vec{A}$$

since $P\vec{C} = 0$.

That is all points on the line are mapped to the same image point $P\vec{A}$, which means that the line must be a ray through the camera center. It follows that \vec{C} is the homogeneous representation of the camera center, since for all choices of \vec{A} the line $\vec{X}(\lambda)$ is a ray through the camera center.

This result is not unexpected since the image point $(0, 0, 0)^T = P\vec{C}$ is not defined, and the camera center is the unique point in space for which the image is undefined. In the case of finite cameras the result may be established directly, since $\vec{C} = (\vec{C}^T, 1)^T$ is clearly the null-vector of $P = \text{KR}[I \mid \vec{C}]$. The result is true even in the case where the first 3×3 sub-matrix M of P is singular. In this singular case, though, the null-vector has the form $\vec{C} = (\vec{d}^T, 0)^T$ where $M\vec{d} = 0$.


The camera center is then a point at infinity. Camera models of this class are discussed in Section 3.3.

A.4 Column vectors

The columns of the projective camera are 3-vectors which have a geometric meaning as particular image points. With the notation that the columns of \mathbf{P} are $\vec{\mathbf{p}}_i$, $i = 1, \dots, 4$, then $\vec{\mathbf{p}}_1, \vec{\mathbf{p}}_2, \vec{\mathbf{p}}_3$ are the vanishing points of the world coordinate X , Y and Z axes respectively. This follows because these points are the images of the axes' directions. For example the x -axis has direction $\vec{\mathbf{D}} = (1, 0, 0, 0)^T$, which is imaged at $\vec{\mathbf{p}}_1 = \mathbf{P}\vec{\mathbf{D}}$. See Figure A.1. The column $\vec{\mathbf{p}}_4$ is the image of the world origin.

A.5 Row vectors

The rows of the projective camera A.4 are 4-vectors which may be interpreted geometrically as particular world planes. These planes are examined next. We introduce the notation that the rows of \mathbf{P} are $\vec{\mathbf{P}}^{iT}$ so that

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{P}}^{1T} \\ \vec{\mathbf{P}}^{2T} \\ \vec{\mathbf{P}}^{3T} \end{bmatrix} \quad (\text{A.4})$$


A.6 The principal plane

The principal plane is the plane through the camera center parallel to the image plane. It consists of the set of points $\vec{\mathbf{X}}$ which are imaged on the line at infinity of the image. Explicitly, $\mathbf{P}\vec{\mathbf{X}} = (x, y, 0)^T$. Thus a point lies on the principal plane of the camera if and only if $\vec{\mathbf{P}}^{3T}\vec{\mathbf{X}} = 0$. In other words, $\vec{\mathbf{P}}^3$ is the vector representing the principal plane of the camera. If $\vec{\mathbf{C}}$ is the camera center, then $\mathbf{P}\vec{\mathbf{C}} = 0$, and so in particular $\vec{\mathbf{P}}^{3T}\vec{\mathbf{C}} = 0$. That is $\vec{\mathbf{C}}$ lies on the principal plane of the camera.

A.7 Axis planes

Consider the set of points $\vec{\mathbf{X}}$ on the plane \mathbf{P}^1 . This set satisfies $\vec{\mathbf{P}}^{1T}\vec{\mathbf{X}} = 0$, and so is imaged at $\mathbf{P}\vec{\mathbf{X}} = (0, y, w)^T$ which are points on the image y -axis. Again it follows from $\mathbf{P}\vec{\mathbf{C}} = 0$ that $\vec{\mathbf{P}}^{1T}\vec{\mathbf{C}} = 0$ and so $\vec{\mathbf{C}}$ lies also on the plane \mathbf{P}^1 . Consequently the plane \mathbf{P}^1 is defined by the camera center and the line $x = 0$ in the image. Similarly the plane \mathbf{P}^2 is defined by the camera center and the line $y = 0$.

Unlike the principal plane \mathbf{P}^3 , the axis planes \mathbf{P}^1 and \mathbf{P}^2 are dependent on the image x - and y -axes, i.e. on the choice of the image coordinate system. Thus they are less tightly coupled to the natural camera geometry than the principal plane. In particular the line of intersection of the planes \mathbf{P}^1 and \mathbf{P}^2 is a line joining the camera center and image origin, i.e. the back-projection of the image origin. This line will not coincide in general with the camera principal axis. The planes arising from \mathbf{P}^i are illustrated in Figure A.2.

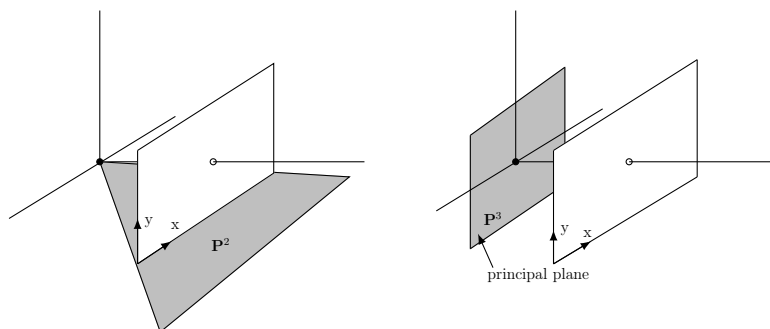


Figure A.2: Two of the three planes defined by the rows of the projection matrix.

The camera center \vec{C} lies on all three planes, and since these planes are distinct (as the P matrix has rank 3) it must lie on their intersection. Algebraically, the condition for the center to lie on all three planes is $P\vec{C} = 0$ which is the original equation for the camera center given above.

A.8 The principal point

The principal axis is the line passing through the camera center \vec{C} , with direction perpendicular to the principal plane \mathbf{P}^3 . The axis intersects the image plane at the principal point. We may determine this point as follows. In general, the normal to a plane $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T$ is the vector $(\pi_1, \pi_2, \pi_3)^T$. This may alternatively be represented by a point $(\pi_1, \pi_2, \pi_3, 0)^T$ on the plane at infinity. In the case of the principal plane \mathbf{P}^3 of the camera, this point is $(p_{31}, p_{32}, p_{33}, 0)^T$, which we denote $\hat{\mathbf{P}}^3$. Projecting that point using the camera matrix P gives the principal point of the camera $P\hat{\mathbf{P}}^3$. Note that only the left hand 3×3 part of $P = [M | \vec{p}_4]$ is involved in this formula. In fact the principal point is computed as $\mathbf{x}_0 = M\vec{m}^3$ where \vec{m}^3 is the third row of M .

A.9 The principal axis vector

Although any point \vec{X} not on the principal plane may be mapped to an image point according to $\vec{x} = P\vec{X}$, in reality only half the points in space, those that lie

in front of the camera, may be seen in an image. Let P be written as $P = [M | \vec{p}_4]$. It has just been seen that the vector \vec{m}^3 points in the direction of the principal axis. We would like to define this vector in such a way that it points in the direction towards the front of the camera (the positive direction). Note however that P is only defined up to sign. This leaves an ambiguity as to whether \vec{m}^3 or $-\vec{m}^3$ points in the positive direction. We now proceed to resolve this ambiguity.

We start by considering coordinates with respect to the camera coordinate frame. According to 3.5, the equation for projection of a 3-D point to a point in the image is given by $\vec{x} = P_{\text{cam}} \vec{X}_{\text{cam}} = K[I | 0] \vec{X}_{\text{cam}}$, where \vec{X}_{cam} is the 3-D point expressed in camera coordinates. In this case observe that the vector $\vec{v} = \det(M) \vec{m}^3 = (0, 0, 1)^T$ points towards the front of the camera in the direction of the principal axis, irrespective of the scaling of P_{cam} . For example, if $P_{\text{cam}} \rightarrow kP_{\text{cam}}$ then $\vec{v} \rightarrow k^4 \vec{v}$ which has the same direction.

If the 3-D point is expressed in world coordinates then $P = kK[R | -R\vec{C}] = [M | \vec{p}_4]$, where $M = kKR$. Since $\det(R) > 0$ the vector $\vec{v} = \det(M) \vec{m}^3$ is again unaffected by scaling. In summary,

- $\vec{v} = \det(M) \vec{m}^3$ is a vector in the direction of the principal axis, directed towards the front of the camera.



Appendix B

Affine Cameras

Consider what happens as we apply a cinematographic technique of tracking back while zooming in, in such a way as to keep objects of interest the same size. This is illustrated in Figure B.1.

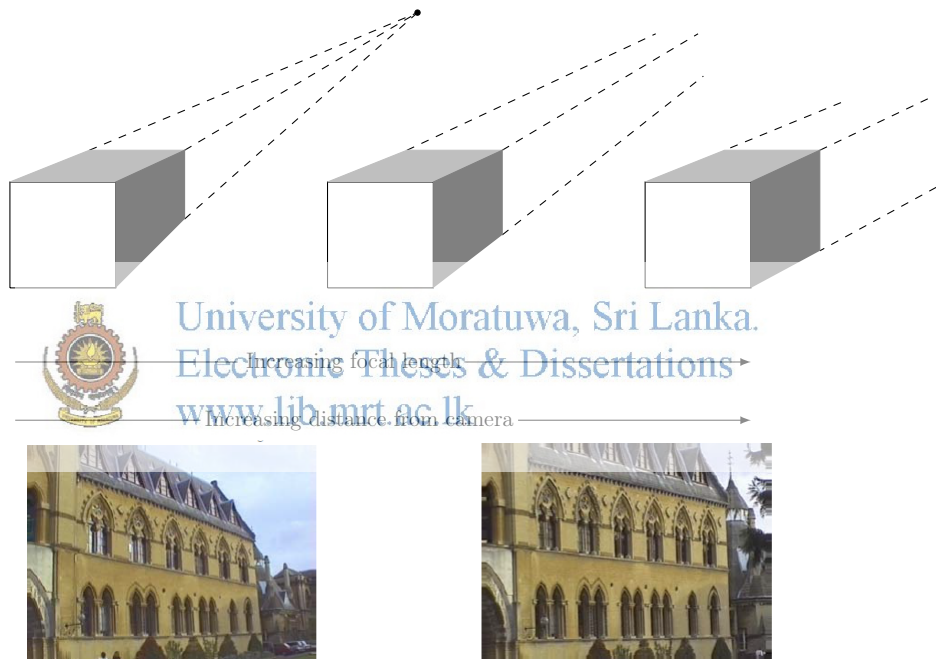


Figure B.1: As the focal length increases and the distance between the camera and object also increases, the image remains the same size but perspective effects diminish. These images are taken from multiple view geometry [41].

We are going to model this process by taking the limit as both the focal length and principal axis distance of the camera from the object increase. In analyzing this technique, we start with a finite projective camera A.3. The camera matrix may be written as

$$P_0 = KR[I \mid -\vec{C}] = K \begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\vec{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\vec{C} \\ \mathbf{r}^{3T} & -\mathbf{r}^{3T}\vec{C} \end{bmatrix} \quad (\text{B.1})$$

where \mathbf{r}^{iT} is the i -th row of the rotation matrix. This camera is located at position \vec{C} and has orientation denoted by matrix R and internal parameters matrix K of the form given in A.2. From Appendix D, the principal ray of the camera is in the direction of the vector \mathbf{r}^3 , and the value $d_0 = -\mathbf{r}^{3T}\vec{C}$ is the distance of the world origin from the camera center in the direction of the principal ray.

Now, we consider what happens if the camera center is moved backwards along the principal ray at unit speed for a time t , so that the center of the camera is moved to $\vec{C} - t\mathbf{r}^3$. Replacing \vec{C} in Equation B.1 by $\vec{C} - t\mathbf{r}^3$ gives the camera matrix at time t :

$$P_t = K \begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}(\vec{C} - t\mathbf{r}^3) \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}(\vec{C} - t\mathbf{r}^3) \\ \mathbf{r}^{3T} & -\mathbf{r}^{3T}(\vec{C} - t\mathbf{r}^3) \end{bmatrix} = K \begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\vec{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\vec{C} \\ \mathbf{r}^{3T} & d_t \end{bmatrix} \quad (\text{B.2})$$

where the terms $\mathbf{r}^{iT}\mathbf{r}^3$ are zero for $i = 1, 2$ because R is a rotation matrix. The scalar $d_t = -\mathbf{r}^{3T}\vec{C} + t$ is the depth of the world origin with respect to the camera center in the direction of the principal ray \mathbf{r}^3 of the camera. Thus

- The effect of tracking along the principal ray is to replace the (3,4) entry of the matrix by the depth d_t of the camera center from the world origin.



University of Moratuwa, Sri Lanka
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Next, we consider zooming such that the camera focal length is increased by a factor k . This magnifies the image by a factor k . It is shown in Appendix D that the effect of zooming by a factor k is to multiply the calibration matrix K on the right by $\text{diag}(k, k, 1)$. Now, we combine the effects of tracking and zooming. We suppose that the magnification factor is $k = d_t/d_0$ so that the image size remains fixed. The resulting camera matrix at time t , derived from B.2, is

$$P_t = K \begin{bmatrix} d_t/d_0 & & \\ & d_t/d_0 & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\vec{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\vec{C} \\ \mathbf{r}^{3T} & d_t \end{bmatrix} = \frac{d_t}{d_0} K \begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\vec{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\vec{C} \\ \mathbf{r}^{3T}d_0/d_t & d_0 \end{bmatrix}$$

and one can ignore the factor d_t/d_0 . When $t = 0$, the camera matrix P_t corresponds with Equation B.1. Now, in the limit as d_t tends to infinity, this matrix becomes

$$P_\infty = \lim_{t \rightarrow \infty} P_t = K \begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T} \vec{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T} \vec{C} \\ \mathbf{0}^T & d_0 \end{bmatrix} \quad (\text{B.3})$$

which is just the original camera matrix B.1 with the first three entries of the last row set to zero. According to the affine camera definition P_∞ is an instance of an affine camera.

B.0.1 Error in employing an affine camera

It may be noted that the image of any point on the plane through the world origin perpendicular to the principal axis direction \mathbf{r}^3 is unchanged by this combined zooming and motion. Indeed, such a point may be written as

$$\vec{X} = \begin{pmatrix} \alpha \mathbf{r}^1 + \beta \mathbf{r}^2 \\ 1 \end{pmatrix}.$$

One then verifies that $P_0 \vec{X} = P_t \vec{X} = P_\infty \vec{X}$ for all t , since $\mathbf{r}^{3T}(\alpha \mathbf{r}^1 + \beta \mathbf{r}^2) = 0$. For points not on this plane the images under P_0 and P_∞ differ, and we will now investigate the extent of this error. Consider a point \vec{X} which is at a perpendicular distance Δ from this plane. The 3D point can be represented as

$$\vec{X} = \begin{pmatrix} \alpha \mathbf{r}^1 + \beta \mathbf{r}^2 + \Delta \mathbf{r}^3 \\ 1 \end{pmatrix}.$$

and is imaged by the cameras P_0 and P_∞ at

$$\vec{\mathbf{x}}_{\text{proj}} = P_0 \vec{X} = K \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ d_0 + \Delta \end{pmatrix} \quad \vec{\mathbf{x}}_{\text{affine}} = P_\infty \vec{X} = K \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ d_0 \end{pmatrix}$$

where $\tilde{x} = \alpha \mathbf{r}^{1T} - \vec{C}$, $\tilde{y} = \beta \mathbf{r}^{2T} - \vec{C}$. Now, writing the calibration matrix as

$$K = \begin{bmatrix} K_{2 \times 2} & \tilde{\mathbf{x}}_0 \\ \tilde{\mathbf{0}}^T & 1 \end{bmatrix},$$

where $K_{2 \times 2}$ is an upper-triangular 2×2 matrix, gives

$$\vec{\mathbf{x}}_{\text{proj}} = \begin{pmatrix} K_{2 \times 2} \tilde{\mathbf{x}} + (d_0 + \Delta) \tilde{\mathbf{x}}_0 \\ d_0 + \Delta \end{pmatrix} \quad \vec{\mathbf{x}}_{\text{affine}} = \begin{pmatrix} K_{2 \times 2} \tilde{\mathbf{x}} + d_0 \tilde{\mathbf{x}}_0 \\ d_0 \end{pmatrix}$$

The image point for P_0 is obtained by de-homogenizing, by dividing by the third element, as $\vec{x}_{\text{proj}} = \tilde{x}_0 + K_{2 \times 2} \tilde{x} / (d_0 + \Delta)$, and for P_∞ the inhomogeneous image point is $\vec{x}_{\text{affine}} = \tilde{x}_0 + K_{2 \times 2} \tilde{x} / d_0$. The relationship between the two points is therefore

$$\vec{x}_{\text{affine}} - \tilde{x}_0 = \frac{d_0 + \Delta}{d_0} \vec{x}_{\text{proj}} - \tilde{x}_0$$

which shows that

- The effect of the affine approximation P_∞ to the true camera matrix P_0 is to move the image of a point \mathbf{X} radially towards or away from the principal point \tilde{x}_0 by a factor equal to $(d_0 + \Delta)/d_0 = 1 + \Delta/d_0$.

This is illustrated in Figure B.2.

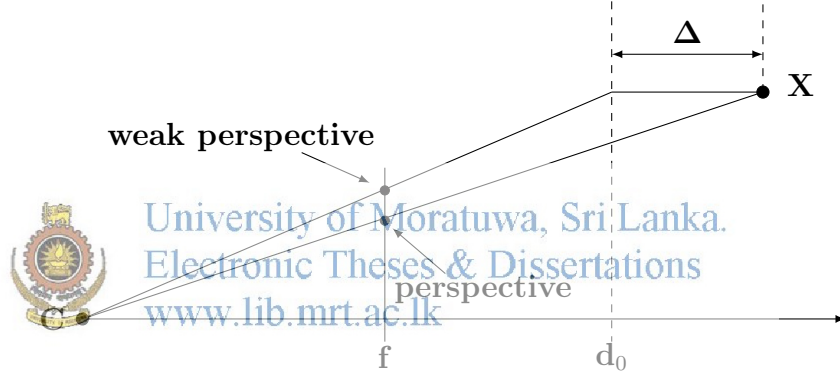


Figure B.2: The action of the weak perspective camera is equivalent to orthographic projection onto a plane (at $Z = d_0$), followed by perspective projection from the plane. The difference between the perspective and weak perspective image point depends both on the distance Δ of the point \mathbf{X} from the plane, and the distance of the point from the principal ray.

B.0.2 Affine imaging conditions

From the expressions for \vec{x}_{proj} and \vec{x}_{affine} we can deduce that

$$\vec{x}_{\text{affine}} - \vec{x}_{\text{proj}} = \frac{\Delta}{d_0} (\vec{x}_{\text{proj}} - \tilde{x}_0) \quad (\text{B.4})$$

which shows that the distance between the true perspective image position and the position obtained using the affine camera approximation P_∞ will be small provided:

- The depth relief (Δ) is small compared to the average depth (d_0), and
- The distance of the point from the principal ray is small.

A general camera matrix of the affine form (P_A), and with no restrictions on its elements, may be decomposed as (Appendix E)

$$P_A = \begin{bmatrix} \alpha_x & s & & \\ & \alpha_y & & \\ & & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ & 1 \end{bmatrix} \quad (\text{B.5})$$

It has eight degrees of freedom, and may be thought of as the parallel projection version of the finite projective camera A.3. In full generality an affine camera has the form

$$P_A = \begin{bmatrix} m_{11} & m_{12} & m_{13} & t_1 \\ m_{21} & m_{22} & m_{23} & t_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{B.6})$$

It has eight degrees of freedom corresponding to the eight non-zero and non-unit matrix elements. We denote the top left 2×3 sub-matrix by $M_{2 \times 3}$. The sole restriction on the affine camera is that $M_{2 \times 3}$ has rank 2. This arises from the requirement that the rank of P is 3.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Appendix C

RQ decomposition

A 3-dimensional Givens rotation is a rotation about one of the three coordinate axes. The three Givens rotations are

$$Q_x = \begin{bmatrix} 1 & & \\ & c & -s \\ & s & c \end{bmatrix} \quad Q_y = \begin{bmatrix} c & & s \\ & 1 & \\ -s & & c \end{bmatrix} \quad Q_z = \begin{bmatrix} c & -s & \\ s & c & \\ & & 1 \end{bmatrix}$$

where $c = \cos(\theta)$ and $s = \sin(\theta)$ for some angle θ and blank entries represent zeros.

Multiplying a 3×3 matrix A on the right by (for instance) Q_z has the effect of leaving the last column of A unchanged, and replacing the first two columns by linear combinations of the original two columns. The angle θ may be chosen so that any given entry in the first two columns becomes zero.

For instance, to set the entry A_{21} to zero, we need to solve the equation $ca_{21} + sa_{22} = 0$. The solution to this is $c = -a_{22}/(a_{22}^2 + a_{21}^2)^{1/2}$ and $s = a_{21}/(a_{22}^2 + a_{21}^2)^{1/2}$. It is required that $c^2 + s^2 = 1$ since $c = \cos \theta$ and $s = \sin \theta$, and the values of c and s given here satisfy that requirement.

The strategy of the RQ algorithm is to clear out the lower half of the matrix one entry at a time by multiplication by Givens rotations. Consider the decomposition of a 3×3 matrix A as $A = RQ$ where R is upper-triangular and Q is a rotation matrix. This may take place in three steps. Each step consists of multiplication on the right by a Givens rotation to set a chosen entry of the matrix A to zero. The sequence of multiplications must be chosen in such a way as not to disturb the entries that have already been set to zero.

An implementation of the RQ decomposition algorithm is given in Table C.1.

Table C.1: Algorithm: RQ decomposition of a 3×3 matrix.

| |
|--|
| <p><u>Objective</u></p> <p>Carry out the RQ decomposition of a 3×3 matrix A using Givens rotations.</p> <p><u>Algorithm</u></p> <p>(i) Multiply by Q_x so as to set A_{32} to zero. (ii) Multiply by Q_y so as to set A_{31} to zero. This multiplication does not change the second column of A, hence A_{32} remains zero. (iii) Multiply by Q_z so as to set A_{21} to zero. The first two columns are replaced by linear combinations of themselves. Thus, A_{31} and A_{32} remain zero.</p> |
|--|

Other sequences of Givens rotations may be chosen to give the same result. As a result of these operations, we find that $AQ_xQ_yQ_z = R$ where R is upper-triangular. Consequently, $A = Q_z^T Q_y^T Q_x^T$, and so $A = RQ$ where $Q = Q_z^T Q_y^T Q_x^T$ is a rotation. In addition, the angles θ_x , θ_y and θ_z associated with the three Givens rotations provide a parametrization of the rotation, by three Euler angles, otherwise known as roll, pitch and yaw angles.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

It should be clear from this description of the decomposition algorithm how similar QR, QL and LQ factorizations may be carried out. Furthermore, the algorithm is easily generalized to higher dimensions.

Appendix D

The importance of the camera center

An object in 3-space and camera center define a set of rays, and an image is obtained by intersecting these rays with a plane. Often this set is referred to as a cone of rays, even though it is not a classical cone. Suppose the cone of rays is intersected by two planes, as shown in Figure D.1, then the two images, I and I' , are clearly related by a perspective map. This means that images obtained with the same camera center may be mapped to one another by a plane projective transformation, in other words they are projectively equivalent and so have the same projective properties. A camera can thus be thought of as a projective imaging device measuring projective properties of the cone of rays with vertex the camera center.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

The result that the two images I and I' are related by a homography will now be derived algebraically to obtain a formula for this homography. Consider two cameras

$$P = KR[I \mid -\tilde{C}], \quad P' = K'R'[I \mid -\tilde{C}]$$

with the same center. Note that since the cameras have a common center there is a simple relation between them, namely $P' = (K'R')(KR)^{-1}P$. It then follows that the images of a 3-space point \vec{X} by the two cameras are related as

$$\vec{x}' = P'\vec{X} = (K'R')(KR)^{-1}P\vec{X} = (K'R')(KR)^{-1}P\vec{x}.$$

That is, the corresponding image points are related by a planar homography (a 3×3 matrix) as $\vec{x}' = H\vec{x}$, where $H = (K'R')(KR)^{-1}$.

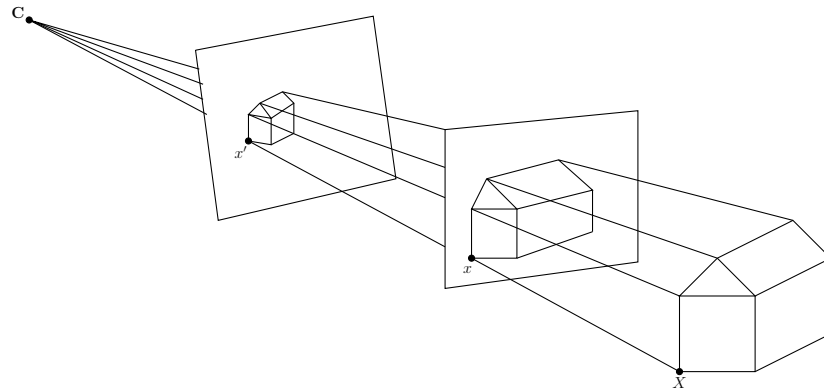


Figure D.1: The cone of rays with vertex the camera center. An image is obtained by intersecting this cone with a plane. A ray between a 3-space point X and the camera center C pierces the planes in the image points x and x' . All such image points are related by a planar homography, $\vec{x}' = H\vec{x}$.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Appendix E

Decomposition of P_∞

The camera matrix obtained by Equation B.3 may be written as

$$P_\infty = \begin{bmatrix} K_{2 \times 2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^T & d_0 \end{bmatrix}$$

where \hat{R} consists of the two first rows of a rotation matrix, $\hat{\mathbf{t}}$ is the vector $(-\mathbf{r}^{1T} \vec{C}, -\mathbf{r}^{2T} \vec{C})^T$, and $\hat{\mathbf{0}}$ the vector $(0, 0)^T$. The 2×2 matrix $K_{2 \times 2}$ is upper-triangular. One quickly verifies that

$$P_\infty = \begin{bmatrix} K_{2 \times 2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^T & d_0 \end{bmatrix} = \begin{bmatrix} d_0^{-1} K_{2 \times 2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^T & d_0 \end{bmatrix}$$

so we may replace $K_{2 \times 2}$ by $d_0^{-1} K_{2 \times 2}$ and assume that $d_0 = 1$. Multiplying out this product gives

$$\begin{aligned} P_\infty &= \begin{bmatrix} K_{2 \times 2} \hat{R} & K_{2 \times 2} \hat{\mathbf{t}} + \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} = \begin{bmatrix} K_{2 \times 2} & \hat{\mathbf{0}}^T \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} + K_{2 \times 2}^{-1} \tilde{\mathbf{x}}_0 \\ \mathbf{0}^T & 1 \end{bmatrix} \\ &= \begin{bmatrix} K_{2 \times 2} & K_{2 \times 2} \hat{\mathbf{t}} + \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{0}}^T \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix}. \end{aligned}$$

Thus, making appropriate substitutions for $\hat{\mathbf{t}}$ or $\tilde{\mathbf{x}}_0$, we can write the affine camera matrix in one of the two forms

$$P_\infty = \begin{bmatrix} K_{2 \times 2} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} = \begin{bmatrix} K_{2 \times 2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix}. \quad (\text{E.1})$$

Consequently, the camera P_∞ can be interpreted in terms of these decompositions in one of two ways, either with $\tilde{\mathbf{x}}_0 = 0$ or with $\hat{\mathbf{t}} = \hat{\mathbf{0}}$. Using the second decomposition of Equation E.1, the image of the world origin is $P_\infty(0, 0, 0, 1)^T = (\tilde{\mathbf{x}}_0, 1)^T$. Consequently, the value of $\tilde{\mathbf{x}}_0$ is dependent on the particular choice of

world coordinates, and hence is not an intrinsic property of the camera itself. This means that the camera matrix P_∞ does not have a principal point. Therefore, it is preferable to use the first decomposition of P_∞ in E.1, and write

$$P_\infty = \begin{bmatrix} K_{2 \times 2} & \hat{\mathbf{o}} \\ \hat{\mathbf{o}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{o}} \\ \hat{\mathbf{o}}^T & 1 \end{bmatrix}$$

where the two matrices represent the internal camera parameters and external camera parameters of P_∞ .



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

References

- [1] O. Javed, Z. Rasheed, K. Shafique, and M. Shah, "Modeling inter-camera space-time and appearance relationships for tracking across non-overlapping views," *Computer Vision and Image Understanding*, vol. 109, no. 2, pp. 146–162, 2008.
- [2] O. Javed, K. Shafique, and M. Shah, "Appearance modeling for tracking in multiple non-overlapping cameras," vol. 2. San Diego, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2005, pp. 26–33.
- [3] O. Javed, Z. Rasheed, K. Shafique, and M. Shah, "Tracking across multiple cameras with disjoint views," vol. 2. Nice, France: Proceedings of the 9th IEEE Conference on Computer Vision, October 2003, pp. 952–957.
- [4] B. C. Matei, H. S. Sawhney, and S. Samarasekera, "Vehicle tracking across nonoverlapping cameras using joint kinematic and appearance features." Colorado Springs, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2011, pp. 3465–3472.
- [5] Y. Sheikh, L. Xin, and M. Shah, "Trajectory association across non-overlapping moving cameras in planar scenes." Minneapolis, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2007, pp. 1–7.
- [6] C. C. Huang, W. Chiu, S. J. Wang, and J. H. Chuang, "Probabilistic modeling of dynamic traffic flow across non-overlapping camera views." Istanbul, Turkey: Proceedings of the 20th IEEE Conference on Pattern Recognition, August 2010, pp. 3332–3335.
- [7] V. Kettner and R. Zabih, "Bayesian multi-camera surveillance," vol. 2. Fort Collins, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1999, pp. 117–123.

- [8] C. Stauffer and K. Tieu, “Automated multi-camera planar tracking correspondence modeling,” vol. 1. Madison, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2003, pp. I-259–I-266.
- [9] Z. Kim, “Real time object tracking based on dynamic feature grouping with background subtraction.” Anchorage, AK: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2008, pp. 1–8.
- [10] Z. Kim and J. Malik, “Fast vehicle detection with probabilistic feature grouping and its application to vehicle tracking,” vol. 1. Nice, France: Proceedings of the 9th IEEE Conference on Computer Vision, October 2003, pp. 524–531.
- [11] T. Collins, Y. Liu, and M. Leordeanu, “Online selection of discriminative tracking features,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1631 – 1643, October 2005.
- [12] N. Buch, J. Orwell, and S. Velastin, “3D extended histogram of oriented gradients (3DHOG) for classification of road users in urban scenes.” London, UK: Proceedings of the British Machine Vision Conference, September 2009, pp. 15.1–15.11.
- [13] J. Shi and C. Tomasi, “Good features to track.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1994, pp. 593–600.
- [14] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, November 2004.
- [15] H. Bay, A. Ess, and L. V. Gool, “Speeded up robust features,” *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, June 2008.
- [16] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” vol. 1. San Diego, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2005, pp. 886 – 893.
- [17] M. Sheikth, Y. and Shah, “Bayesian modeling of dynamic scenes for object detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1778–1792, November 2005.

- [18] J. Yang, Y. Wang, A. Sowmya, and Z. Li, "Vehicle detection and tracking with low-angle cameras." Hong Kong: Proceedings of the 17th IEEE International Conference on Image Processing, September 2010, pp. 685–688.
- [19] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Computing Surveys*, vol. 38, no. 4, December 2006, article 13.
- [20] N. Buch, A. Velastin, and J. Orwell, "A review of computer vision techniques for the analysis of urban traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 920–939, September 2011.
- [21] V. Kastrinaki, M. Zervakis, and K. Kalaizakis, "A survey of video processing techniques for traffic applications," *Image and Vision Computing*, vol. 21, no. 4, pp. 359–381, 2003.
- [22] M. Valera and S. A. Valastin, "Intelligent distributed surveillance systems: a review," *IEE Proceedings of Vision, Image and Signal Processing*, vol. 152, no. 2, pp. 192–204, April 2005.
- [23] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 694–711, May 2006.
- [24] D. J. Jayamanne, J. Samarawickrama, and R. Rodrigo, "Appearance based tracking with background subtraction." Proceedings of the IEEE Conference on Computer Science and Education, April 2013, pp. 643–649.
- [25] S. Birchfield, "Derivation of Kanade-Lucas-Tomasi tracking equation," January 1997.
- [26] H. Schweitzer, J. W. Bell, and f. Wu, "Very fast template matching." Copenhagen, Denmark: Proceedings of the 7th European Conference on Computer Vision, May 2002, pp. 358–372.
- [27] J. Yang, Y. Wang, A. Sowmya, Z. Li, B. Zhang, and J. Xu, "Ieee workshop on applications of computer vision (wacv)," Kona, HI, January 2011, pp. 382–388.
- [28] D. Koller, K. Daniilidis, and H. H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *International Journal of Computer Vision*, vol. 10, no. 3, pp. 257–281, June 1993.

- [29] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell, "Towards robust automatic traffic scene analysis in real-time," vol. 1. Jerusalem, Israel: Proceedings of the 12th IAPR International Conference on Pattern Recognition, October 1994, pp. 126–131.
- [30] D. Koller, J. Weber, and J. Malik, "Robust multiple car tracking with occlusion reasoning." Stockholm, Sweden: Proceedings of the 3rd European Conference on Computer Vision, May 1994, pp. 189–196.
- [31] B. Coifman, D. Beymer, P. McLauchlan, and J. Malik, "A real-time computer vision system for vehicle tracking and traffic surveillance," *Transportation Research Part C: Emerging Technologies*, vol. 6, no. 4, pp. 271–288, August 1998.
- [32] N. K. Kanhere, S. J. Pundlik, and S. T. Birchfield, "Vehicle segmentation and tracking from a low-angle off-axis camera," vol. 2. San Diego, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2005, pp. 1152–1157.
- [33] N. K. Kanhere and S. T. Birchfield, "Real-time incremental segmentation and tracking of vehicles at low camera angles using stable features," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 148–160, March 2008.
- [34] B. Tamersoy and J. Aggarwal, "Robust vehicle detection for tracking in highway surveillance videos using unsupervised learning." Genova, Italy: Proceedings of the 6th IEEE International Conference on Advanced Video and Signal Based Surveillance, September 2009, pp. 529–534.
- [35] J. Shi and J. Malik, "Normalized cuts and image segmentation." San Juan, Puerto Rico: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1997, pp. 731–737.
- [36] G. Jun, A. J. K., and M. Głokmen, "Tracking and segmentation of highway vehicles in cluttered and crowded scenes." Copper Mountain, USA: IEEE Workshop on Applications of Computer Vision, January 2008, pp. 1–6.
- [37] M. Brulin, H. Nicolas, and C. Maillet, "Video surveillance traffic analysis using scene geometry." Singapore: Proceedings of the 4th Pacific-Rim Symposium on Image and Video Technology, November 2010, pp. 450–455.

- [38] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer, 2011.
- [39] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.
- [40] R. O. Duta, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley, 2001.
- [41] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [42] J. Y. Bouguet, “Camera calibration toolbox for matlab,” http://www.vision.caltech.edu/bouguetj/calib_doc/, 2008.
- [43] G. Bradski and A. Kaehler, *Learning OpenCV*. O’ Reilly Media Inc., 2008.
- [44] R. Laganière, *OpenCV 2 Computer Vision Application Programming Cookbook*. Packt Publishing Ltd, 2011.
- [45] W. Savich, *Absolute C++*. Addison Wesley, 2002.
- [46] Y. Fan, H. Yang, S. Zheng, and H. Su, “Geometric motion flow(gmf): A new feature,” Hefei, China: Proceedings of the 6th International Conference on Image and Graphics, August 2011, pp. 726–730.
- [47] C. C. C. Pang, W. Lam, and N. Yung, “A method for vehicle count in the presence of multiple-vehicle occlusions in traffic images,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 3, pp. 441–459, September 2007.
- [48] J. Yang, Y. Wang, A. Sowmya, Z. Li, B. Zhang, and J. Xu, “Feature fusion for vehicle detection and tracking with low-angle cameras.” Kona, Hawaii: IEEE Workshop on Applications of Computer Vision, January 2011, pp. 382–388.
- [49] R. C. Gonzalez and W. R. E., *Digital Image Processing*. Prentice Hall, 2007.
- [50] R. C. Gonzalez, W. R. E., and S. L. Eddins, *Digital Image Processing Using MATLAB*. Gatesmark Publishing, 2009.
- [51] C. Bishop, *Pattern recognition and machine learning*. Springer Science Business Media, LLC, 2006.

- [52] R. Xu and D. C. Wunsch, "Survey of clustering algorithms," *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, May 2005.
- [53] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," vol. 2. Fort Collins, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1999.
- [54] S. Brutzer, B. Hoferlin, and G. Heidemann, "Evaluation of background subtraction techniques for video surveillance." Rhode Island, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2011, pp. 1937–1944.
- [55] M. Piccardi, "Background subtraction techniques: a review," vol. 4. Hague, Netherlands: Proceedings of the IEEE Conference on Systems, Man and Cybernetics, October 2004, pp. 3099–3104.
- [56] D. H. Parks and S. S. Fels, "Evaluation of background subtraction algorithms with post-processing." Santa Fe, USA: Proceedings of the 5th IEEE Conference on Advanced Video and Signal Based Surveillance, September 2008, pp. 192–199.
- [57] Y. Benezeth, P. McJannet, E. Fumier, H. Laurent, and C. Rosenberger, "Review and evaluation of commonly implemented background subtraction algorithms." Tampa, USA: Proceedings of the 19th International Conference on Pattern Recognition, December 2008, pp. 1–4.
- [58] J. Batista, P. Peixoto, C. Fernandes, and M. Ribeiro, "A dual-stage robust vehicle detection and tracking for real-time traffic monitoring." Toronto, Canada: Proceedings of the IEEE Intelligent Transportation Systems Conference, September 2006, pp. 528–535.
- [59] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting objects, shadows and ghosts in video streams by exploiting color and motion information." Palermo, Italy: Proceedings of the 11th International Conference on Image Analysis and Processing, September 2001, pp. 360–365.
- [60] —, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1337–1342, October 2003.

- [61] T. S. F. Haines and T. Xiang, "Background subtraction with dirichlet processes." Florence, Italy: Proceedings of the 12th European Conference on Computer Vision, October 2012, pp. 99–113.
- [62] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, November 2001.
- [63] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.
- [64] S. Oh, A. Hoogs, A. Perera, N. Cuntoor, C. Chen, C., T. Lee, J., S. Mukherjee, K. Aggarwal, J., H. Lee, L. Davis, E. Swears, X. Wang, Q. Ji, K. Reddy, M. Shah, C. Vondrick, H. Pirsiavash, D. Ramanan, J. Yuen, A. Torralba, B. Song, A. Fong, R. Chowdhury, A., and M. Desai, "A large-scale benchmark dataset for event recognition in surveillance video." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 2011, pp. 3153–3160.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk