Robot Perception
Assignment 06

Anand Ajmera
José Carlos Mayoral Baños

May 17, 2016

Summary Last Lecture

Cameras

- Continues....
- General Cameras

- Matrix P representations:
- Columns.. represents the vanishing points, in other world the world axis directions. Rows... axis planes (x=0 and y=0) and principle plane.
- General projective camera (P) maps x = PX
- Matrix P is rank 3, with just 1D null space.
- Null space is a 4D vector which shows the camera center at homogeneous coordinates.
- If the null space is finite then the camera is finite, other wise is a infinite camera.
- The principal plane for finite centers a 4 vector abcd represents the equation ax+by+cz+d = 0, remembering that points in the principal plane are the ones that are mapped to the infinity.
- Principal axis is a normal of the principal plane.
- Matrix P divides itself in blocks M—p4.
- The last row of M shows the principal axis direction.
- To find a backprojection from the image poitn to the projective space the pseudoinverse is needed.

- Having the matrix P and the projective of a point X to x= PX we can compute the depth since PC = 0 being C the center of the camera.
- w= m3(X-X) where m3 is the direction of the principal axis.
- Normalizing M and m3 then w is the depth of X from C in the direction of principal axis.
- SVD can be used to obtain intrinsic and extrinsic parameters.
- Last column of V s the null space or C.
- Other way to analyze P is taking M and use QR to find KR where K is the calibration matrix and R represents the homogeneous transformation.
- K matrix considers skew and focal length.
- For Real world homographies are euclidean, general are used for generalization.
- Infinity cameras has a singular M.
- For calibration, given correspondences are used to calculate P.
- Restricted camera estimation problem gives correspondences and some knowledge of P.
- In vision, true projective camera is assumed but it doesn't consider lens distortion.
- DLT or Gold standard can be used to compute P.
- Lens distortion is solved considering it as radial.

• Multiple View Geometry.

- Represent geometric relations between two views of the same plane.
- Solver 3 problems.
- Knowing matrix P and x how will be the correspondence point in the other plane.
- Given both corresponding point, calculates P and P'
- Give P, P' and both correspondencies, calculate 3D position of the points.
- If both views are just a translation are called auto-epipolar.
- Epipolar Geometry.
- Projective geometry of two views that just depends on the camera's internal parameters and relative poses of the cameras.

- A matrix F captures the geometry of the epipolar geometry.
- F computing is similar to homography matrix.
- With F, P and P' camera properties matrix can be deduced.
- Baseline connects both camera centers.
- Epipolar plane is created with the baseline and the correspondence points.
- Intersection of baseline and image planes are called Epipoles.
- At rotating points the whole plane rotates according the baseline.
- Epipolar line is the intersection of an epipolar plane with the image plane.
- For correspondence points between two image we only need to look for the correspondence along the epipolar line.
- All epipolar lines intersect at the epipole.

Summary

Section 6.2 The projective camera

A general projective camera P maps wold point to image points. We can find geometric entities which defines the camera model.

Camera anatomy

General projective P is discomposed in M and p4 matrices, with M as a non singular matrix being a finite camera. Matrix P has a 1D null space due to it's rank 3 and 4 columns, this null space shows the camera center vector. This vector line contains also other points that relies on the vector. All this points A points to the same image point. But from all this points, the image is defined expect for the camera center. But for infinite cameras the last element of P is 0 and shows that the camera center is at the infinity.

Columns of P represents the vanishing points of the world coordinate X Y and Z axes. Because these points are images of the axes directions. Rows the projective camera can be interpreted geometrically as particular world planes within the last row represents the principal plane.

Principal plane is the plan through the camera center parallel to the image plane. Axis planes are imaged as PX = (0,y,w) and PX = (x,0,w).

Intersection line of the planes P 1 and P 2 is a line joining the camera center and image origin. In general, this line will not coincide the camera principal axis. C lies on all three planes, and since these planes are distinct, C lies on intersections.

Principal point is the line passing through the camera center C with direction perpendicular to the principal plane. The axis intersects image plane and principal point. Principal axis vector may be mapped to an image point according x = PX.

Action of a projective camera on points

Forward projection: General projective camera maps a point in space X to an image point. Back-projection of point to rays: Given a point x in an image, we next determine the set of point in space that map to this point. These are the camera center C and the point P+x, where P+ is the pseudo-inverse of P. Point P+x lies on the ray because it projects to x, since P(P+x) = Ix = x.

In the case of finite cameras an alternative expression can be developed. An image point x back-projects to a ray intersecting the plane at infinity at the point D.

Depth of points

We consider the distance a point lies in front of or behind the principal plane of the camera. Any camera matrix may be normalized by multiplying it by an appropiate factor. Avoid having always to deal with normalized camera matrices.

Decomposition of the camera matrix

Let P be a camera matrix representing a general projective camera. To find the camera orientation and the internal parameters of the camera. Camera center can be obtained by SVD of P, as well as the camera orientation and internal parameters for finite camera an QR method can be implemented obtaining matrixes K and R where K in the calibration matrix and R gives the orientation.

For skew different to zero means that the axis x an y are not perpendicular which is not likely that is a result of taking an image of an image. The most severe distortion comes from this picture's picture process is a planar homography. The original camera is represented by P then the camera of picture's picture is a HP where H is homography. In this case QR results matrixes are not KR.

A common practice in measuring image coordinates is that the y-coordinate increases in the downwards direction. A recommended practice is to negate the coordinate of the image point so the coordinate system again becomes right handed. Decomposition of this camera matrix with K is still possible with a x and a y positive. The difference is that R now represents the orientation of the camera with respect to the negative z-axis.

Euclidean vs projective spaces

Until now a euclidean space has been considered. From now, use a projective coordinate frame. With a world coordinate frame is projective a transformation between the camera and world coordinate frame is reached. The interpretation of P 3 as the principal plane requires at least affine frames.

Section 7.1 Basic Equations

Assuming correspondences between 3D point to 2D points, camera matrix P must be found. This problem has been seen before but with a different dimension. But still be the same due to the linear dependence we can use 2D projective transformation approach.

Matrix P also has 12 entries but the scale is ignored. Five and a half correspondence points are needed to solve P, Ap=0. With minimum correspondences, solution is exact. A has only 1D null space.

If the data is not exact for overdetermined solution with more correspondence points. The estimation of a homography is gotten by minimizing the residual. The complete DLT algorithm for computation of P proceeds in the same manner as that.

If data is close to a degenerate configuration then a poor estimate for P is obtained due to a unique solution for P cannot be obtained. The camera and points lie on a twisted cubic and the point all lie on the union of a plane and a single straight line containing the camera center.

Normalizing data also applies to 3D homography estimation case translating the points with the distance to the origin limited to the square root of 3.

In line correspondences, it's possible to extend DLT to take in account line correspondences as well which are represented by two points throigh the line passes.

Section 7.2 Geometric Error

Supposing that the world points are given so there are more accurate that the measured image points. So the geometric error is the sum of the error between measured points and estimated points. IF the errors are gaussian then the solution must be minimized calling this Maximum Likelihood.

DLT solution, or a minimal solution, may be used as a starting point for the iterative minimization as the same of 2D homohraphy case.

Perhaps, the world points are not exactly so a minimize so a 3D geometric or image geometric error minimization has to be done between a closest reachable point in space and estimated points. Therefore if error at world and image errors are considered, has to be mimized as well by Mahalanobis distance with respect to the known error covariance matrices for each of the measurements.

Geometric interpretation of algebraic error

In an ideal case that all points are normalized, the depth from the camera in the direction along principal ray are taken to estimate p3. Also depth (Zi) are used to compute the depth The distance d(X i ,X' i) is the correction that needs to be made to the measured 3D points in order to correspond precisely with the measured image points x.

Estimation of an affine camera.

The methods developed above for the projective cameras can be applied directly to affine cameras. In the case of computing 2D affine transformations for affine cameras, algebraic error and geometric image error are equal. This is because affinity represents a scale so both principal axis has the same direction.

Section 7.3 Restricted camera estimation

DLT computes a general projective camera matrix P to map 3D to 2D point correspondences. Calibration matrix contains an upper matrix with non-zero entries. There are some common assumption to find the best-fit camera matrix such as:

1. Skew = 0.
2. Pixels are square.
3. Principal point is known.
4. Compete calibration matrix K is known.

To minimize geometric error, one selects a set of parameters that characterize the camera matrix to be computed. The geometric error may then be minimized with respect to the set of parameters using iterative minimization.

It is possible to minimize algebraic error instead, in which case the iterative minimization problem becomes much smaller. Given a set of n world to image correspondences, $Xj < -Xj$, the problem of finding a constrained camera matrix P that minimizes the sum of algebraic distances $J2i^{a}/p(xi; PXj)$ 2 reduces to the minimization of a function $IR9 > IR, 12$, independent of the number n of correspondences.

Knowing internal parameters, remaining parameters are the position and orientation of the camera. This is the "exterior orientation" problem, which is important in the analysis of calibrated systems. There are six parameters that must be determined, three for the orientation and three for the position.

Camera Matrix P

Build the camera matrix P for your smartphone camera. Use MATLAB (or using the CALTECH camera calibration toolbox or openCV with phyton wrapper).

We chose a Asus Tablet Camera to create the calibration matrix using MATLAB toolbox.

The undistorted image can been shown using MATLAB function imshow().

The calibration parameters are shown in the "out.pdf" file.