회전과 줄을 하는 카메라의 Self-Calibration

서용덕* 흥기상
포항공과대학교 전자전기공학과 영상처리연구실

Self-Calibration of a Rotating and Zooming Camera

Yongduck Seo Ki Sang Hong
MVIP Lab., EE, Depl.
Pohang University of Science and Technology (POSTECH)
dragon@indy.postech.ac.kr hongka@vision.postech.ac.kr

요 약
이 논문에서는 회전과 줄을 하는 카메라의 내부변수를 3차원 투영 없이 추정하는 방법을 제안한다. 먼저, 카메라의 skew을 0으로 가정하면 카메라의 내부변수가 매 영상씩 독립적으로 결정되며, 그 값을 구할 수 있다는 것을 이론적으로 보인다. 이에 구해지는 회전 환경은 기준 좌표계를 설정하는데 따라 달라질 뿐이다. 카메라 모델은 확률.mx는 아웃, 회전이 두 영상 사이에 투영변환선이 존재한다는 것을 이용한 것이다. 가장 일반적인 경우, 즉 카메라의 skew을 0으로 가정하면, 카메라 내부변수는 계산하기 쉬워 대부분의 경우에 투영변환선이 필요하다. 이는, 카메라의 principal point가 시간에 따라 변화하지 않는다고 가정하면 두 가지의 투영변환선이 필요하며, 가장 일반적인 카메라 모델의 경우 principal point와 aspect ratio가 변하지 않으면 단지 한 개의 투영변환선이 필요하다. 허용되는 실제 영상 데이터를 이용하여 계산하는 알고리즘을 설명하였다.

1. Introduction

Recently, there have been lots of researches for calibrating a camera based only on matches of multiple images. They reported algorithms of auto-calibration for fixed internal camera parameters [5, 2, 8] Applying the techniques of Projective Geometry, they showed that it is possible to compute the five internal camera parameters when they are fixed for all the views. When camera parameters are varying from image to image, then under the assumption that at least one of five internal parameters is known, auto-calibration is possible[3, 4, 6] All these auto-calibration methods require that the views be taken at different viewpoints. That is translation is not zero.

Harley proposed a self-calibration algorithm given matches of images taken by a rotating camera whose internal parameters are fixed[1] One limitation of the work is that the algorithm cannot be applied when the images are taken by a zooming or auto-focusing camera, which is common in video images of sports games like soccer or American football. In this paper, we propose a method to auto-calibrate such a rotating and zooming camera so that 3D information can be extracted for future analysis.

2. Self-Calibration is Possible

In this paper we consider a set of rotating cameras with camera matrices \( P_k = K_k[R_k|0] \), where \( K_k \) is the camera calibration matrix of zero skew defined by

\[
K_k = \begin{bmatrix}
\alpha_k & 0 & x_k \\
0 & \beta_k & y_k \\
0 & 0 & 1
\end{bmatrix}
\]

The parameters in \( K_k \), the intrinsic parameters, represent the properties of the image formation system: \( \beta_k \) represents focal length, \( \gamma_k = \alpha_k/\beta_k \) represents the aspect ratio and \( (x_k, y_k) \) is the principal point. The parameter \( R_k \) is the rotation of the \( k \)-th camera with respect to the 0-th. Also, the inter-image homography can be computed from image matches and satisfy the relationship \( u_k = H_k u_0 \) where \( u_k \) and \( u_0 \) are matching points.

Using the inter-image homographies \( H_k \)'s computed from image matches, we can find camera matrices \( P_k = K_k[R_k|0] \), \( k = 0, \ldots, N \), that satisfy the relationship \( H_k = P_k P_0^{-1} = K_k K_k^0 \). Notice that given one such sequence of camera matrices \( P_k \), \( k = 0, \ldots, N \), \( P_k Q \) may also be a possible choice of camera matrices, where \( Q \) is a nonnegative 3 x 3 matrix, because they also produce the same inter-image homographies.

Now we need a lemma to go further [4, 6].

Lemma 1 A camera matrix \( P = KR = [p_1 \ p_2 \ p_3] \) represents a zero-skew camera if and only if

\[
(p_1 \times p_2) \cdot (p_2 \times p_3) = 0 \tag{2}
\]

Due to this lemma, the projective transformation \( Q_{3\times3} \) cannot be arbitrary because every camera matrix should satisfy the constraint equation (2).
Now it remains to show that given a sequence of camera matrices \( P_k, k = 1, \ldots, N \), which 1) solves the inter-image transformation problem and 2) represents zero-skew cameras, then the only possible transformations \( Q_{3 \times 3} \) that preserve the zero-skew camera condition (equation (2)) are the orthogonal transformations.

Denote by \( M_P \) the manifold of all \( 3 \times 3 \) camera matrices defined up to scale. Denote by \( M_{3 \times 3} \) the manifold of all camera matrices that represent zero-skew cameras. Denote the group of all projective transformations, represented by \( 3 \times 3 \) matrices, by \( G_P \). Finally, denote by \( G_{3 \times 3} \) the group of transformations that preserve the property in Lemma 1, and the group of orthogonal transformations by \( G_O \):

\[
G_{3 \times 3} = \{ Q \in G_P | (P \in M_{3 \times 3}) \Rightarrow PQ \in M_{3 \times 3} \}
\]

\[
G_O = \{ \lambda Q | Q Q^T = I, 0 \neq \lambda \in R \}
\]

It is clear that the group of orthogonal transformations is contained in \( G_{3 \times 3} \). If \( G_{3 \times 3} = G_O \), then it is possible to calibrate cameras uniquely up to orthogonal transformation.

**Theorem 1** Let \( G_{3 \times 3} \) denote the class of transformations that preserve the zero-skew camera condition and \( G_O \) the group of orthogonal transformations. Then

\[
G_{3 \times 3} = G_O.
\]

**Proof:** It is clear that \( G_O \subseteq G_{3 \times 3} \). Now we show that \( G_O \supseteq G_{3 \times 3} \). Assume that \( P \) represents a zero-skew camera, \( Q \) a projective transformation in \( G_{3 \times 3} \). Then, from the definition, \( PQ = KRQ \) can be re-written in the form of \( K' R' \) where \( K' \) is a zero-skew calibration matrix and \( R' \) is an orthogonal matrix. Also \( UQV \) has this property for every pair of orthogonal matrices \( U \) and \( V \), since

\[
KRUVQ = K'R'UV = K'R'
\]

where \( R'' \) and \( R''' \) denote orthogonal matrices. Now, using singular value decomposition of \( Q \) we may write

\[
D = UQV = \begin{bmatrix} d_1 & d_2 & d_3 \end{bmatrix}
\]

Suppose that the rotation matrix \( R \) is given by \( \theta \) degree rotation about x-axis and by \( \phi \) degrees about y-axis (note that the rotation \( R \) is arbitrary)

\[
R = \begin{bmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos \phi \cos \theta & \sin \phi \sin \theta & -\cos \phi \cos \theta \\ \sin \phi \cos \theta & \cos \phi \sin \theta & \cos \phi \cos \theta \\ -\sin \phi \sin \theta & \cos \phi \cos \theta & \cos \phi \cos \theta \end{bmatrix},
\]

we have

\[
RD = \begin{bmatrix} d_1 \cos \phi & d_2 \sin \phi \sin \phi & d_3 \cos \phi \sin \phi \\ 0 & d_2 \cos \phi & d_3 \sin \phi \\ d_1 \sin \phi & -d_2 \sin \phi \cos \phi & d_3 \cos \phi \cos \phi \end{bmatrix}.
\]

Now, according to Lemma 1, \( RD = [r_1, r_2, r_3]^T \) is a zero-skew calibration matrix if and only if \( (r_1 \times r_2) \cdot (r_2 \times r_3) = 0 \). After some calculation we have

\[
r_1 \times r_2 = \begin{bmatrix} 0, -d_1 d_3 \cos \theta, -d_1 d_2 \sin \theta \end{bmatrix}^T
\]

\[
r_2 \times r_3 = \begin{bmatrix} d_2 d_3 \cos \phi, d_1 d_3 \sin \theta \sin \phi, -d_1 d_2 \cos \theta \sin \phi \end{bmatrix}^T
\]

and

\[
(r_1 \times r_2) \cdot (r_2 \times r_3)
\]

\[
= -d_1^2 d_2^2 \cos \phi \sin \theta \sin \phi + d_1^2 d_3^2 \cos \theta \sin \phi \sin \phi
\]

\[
= d_1^2 \cos \theta \sin \theta \sin \phi(-d_3^2 + d_3^2)
\]

\[
= 0.
\]

That is, \( RD \) is a zero-skew calibration matrix if and only if \( d_3 = d_3 \). Permutation of the singular values yields \( d_1 = d_2 = d_3 \). Thus, all singular values of \( Q \) are equal, which means that \( Q \) is an orthogonal matrix.

**3. Estimation Method**

From the equation (1), we have \( H_k K_0 K_0^T H_k^T = K_0 K_0^T \), and if we know the principal points \( (x_k, y_k) \), other calibration parameters \( (\alpha_k, \beta_k) \) can be computed using a linear equations. In other words, the scale parameters \( (\alpha_k, \beta_k) \) may be parametrized by the principal points, and given principal points the scale parameters are linearly computed. Now we define a non-linear error function to find the optimal calibration parameters including the principal points. Using the relationship

\[
R_k = \frac{K_0^{-1} H_k K_0}{\det(K_0^{-1} H_k K_0)}^{1/2}
\]

we minimize the following error function

\[
E = \sum_{k=1}^{N} \left( \| R_k R_k^T - I_{3 \times 3} \|^2_F + \| R_k^T R_k - I_{3 \times 3} \|^2_F \right)
\]

Notice that \( E \) is a function of principal points. Since the principal points are around image center, a search window may be chosen around the image center, and the algorithm proposed is:

1. Set principal points \( x_k \leftarrow \tilde{x}_k \) and \( y_k \leftarrow \tilde{y}_k \) for \( k = 0, \ldots, N \)
2. Compute \( \alpha_k \) and \( \beta_k \) for \( k = 0, \ldots, N \)
3. Compute the error \( E \) defined by equation (4).
4. If \( E \) is smaller than the previous one, record the calibration parameters
5. Repeat step 1-4 for searching area
6. The optimal calibration parameters are the recorded ones.

**4. Experiments**

Here, we show results of our algorithm using two views. Assuming that the aspect ratio is 1 and the principal points are fixed, auto-calibration can be done using only two views. Table 1 shows the calibration result of 100 runs with 2 image matches
for various image noise. Since the noise is added to each of image coordinates, the actual RMS error is $\sqrt{2}$ times the indicated value $\sigma$. Note that the principal point is the most sensitive to input image noise. On the contrary, rotation angles are less sensitive to input noise.

Figure 1 shows two video frames of a soccer game. Notice that there are scale change due to zooming as well as rotation. Inter-image homography is estimated by direct iterative error minimization method [7] where initial parameters are obtained using matches of lines and points, and the calibration result is: $f_0 = 1145.9$, $f_1 = 1376.5$ and $(x, y) = (324.5, 181.5)$. Computed rotation angles for the three axes are $(-3.75^\circ, -10^\circ, 0^\circ)$.

5. Conclusion

We showed that auto-calibration of a rotating and zooming camera without 3D pattern is unique up to orthogonal transformation and implemented and tested the algorithm for synthetic and real data. This algorithm is important for the applications like 3D reasoning from monocular rotating camera in sports games or video re-generation of a scene based on image mosaic.

참고문헌


