회전과 줌을 하는 카메라의 Self-Calibration

서 용덕° 홍 기상
포항공과대학교 전자전기공학과 영상처리연구실

Self-Calibration of a Rotating and Zooming Camera

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

요 약

이 논문에서는 회전과 줌을 하는 카메라의 내부변수를 3차원 패턴 없이 주어진 영상만을 이용하여 구하는 방법을 제안한다. 먼저, 카메라의 skew를 0 으로 가정하면 카메라의 내부변수가 매 영상취득 시점마다 바뀌어도 그 값들을 구할수 있다는 것을 이론적으로 보인다. 이때 구해지는 회전 행렬은 기준 좌표계를 설정하는데 따라 달라질 뿐이다. 카메라 보정은 획득되어진 영상 사이의 투영변환행렬을 분석하여 얻어지며, 이는 회전하는 카메라에서 얻어진 두 영상 사이에는 투영변환행렬이 존재한다는 것을 이용한 것이다 가장 일반적인 경우, 즉 카메라의 skew를 0으로 가정하면, 카메라 내부변수를 계산하기 위하여 모두 네 개의 투영변환행렬이 필요하다. 또, 카메라의 principal point가 시간에 따라 변화하지 않는다고 가정하면 두 개의 투영변환행렬이 필요하며, 가장 단순한 카메라 모델의 경우 principal point 와 aspect ratio 가 변화하지 않으면 단지 한 개의 투영거리변환이 필요하다. 합성 데이타와 실제 영상 데이타를 이용하여 제안하는 알고리듬을 시험하였다

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

Hartley 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 $\mathbf{P}_k = \mathbf{K}_k[\mathbf{R}_k|0]$, where $\mathbf{K}_k$ is the camera calibration matrix of zero skew defined by

$$\mathbf{K}_k = \begin{bmatrix} \alpha_k & 0 & x_k \\ & \beta_k & y_k \\ & & 1 \end{bmatrix}$$

The parameters in $\mathbf{K}_k$, the *intrinsic parameters*, represent the properties of the image formation system: β_k represents *focal length*, $\gamma_k = \alpha_k/\beta_k$ represents the *aspect ratio* and (x_k, y_k) is called the *principal point*

Note that given a set of images $\mathcal{I}_0, \ldots, \mathcal{I}_N$ taken from the same location by cameras with the calibration matrices $\mathbf{K}_k$, then there exist 2D projective transformations $\mathbf{H}_k$, taking image $\mathcal{I}_0$ to image $\mathcal{I}_k$, whose matrices are of the form.

$$\mathbf{H}_k = \mathbf{K}_k\mathbf{R}_k\mathbf{K}_0^{-1} \quad (1)$$

where $\mathbf{R}_k$ represents 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 $\mathbf{u}_k = \mathbf{H}_k\mathbf{u}_0$ where $\mathbf{u}_k$ and $\mathbf{u}_0$ are matching points.

Using the inter-image homographies $\mathbf{H}_k$'s computed from image matches, we can find camera matrices $\mathbf{P}_k = \mathbf{K}_k\mathbf{R}_k$, $k = 0, ..., N$, that satisfy the relationship $\mathbf{H}_k = \mathbf{P}_k\mathbf{P}_0^{-1} = \mathbf{K}_k\mathbf{R}_k\mathbf{K}_0^{-1}$ Notice that given one such sequence of camera matrices $\mathbf{P}_k, k = 0, ..., N$, $\mathbf{P}_k\mathbf{Q}$ may be also a possible choice of camera matrices, where $\mathbf{Q}$ is a nonsingular 3×3 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* $\mathbf{P} = \mathbf{KR} = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3]^T$ *represents a zero-skew camera if and only if*

$$(\mathbf{p}_1 \times \mathbf{p}_3) \cdot (\mathbf{p}_2 \times \mathbf{p}_3) = 0 \quad (2)$$

Due to this lemma, the projective transformation $\mathbf{Q}_{3\times3}$ can not be arbitrary because every camera matrix should satisfy the constraint equation (2).

Now it remains to show that given a sequence of camera matrices $\mathbf{P}_k$, $k = 1, ..., N$, which 1) solves the inter-image transformation problem and 2) represents zero-skew cameras, then the only possible transformations $\mathbf{Q}_{3\times 3}$ that preserve the zero-skew camera condition (equation (2)) are the orthogonal transformations.

Denote by $\mathcal{M}_\mathbf{P}$ the manifold of all 3×3 camera matrices defined up to scale. Denote by $\mathcal{M}_{zs}$ the manifold of all camera matrices that represent zero-skew cameras. Denote the group of all projective transformations, represented by 3×3 matrices, by $\mathcal{G}_\mathbf{P}$. Finally, denote by $\mathcal{G}_{zs}$ the group of transformations that preserve the property in Lemma 1, and the group of orthogonal transformations by $\mathcal{G}_\mathbf{O}$:

$$\mathcal{G}_{zs} = \{\mathbf{Q} \in \mathcal{G}_\mathbf{P} | (\mathbf{P} \in \mathcal{M}_{zs}) \Rightarrow \mathbf{PQ} \in \mathcal{M}_{zs}\}$$
$$\mathcal{G}_\mathbf{O} = \{\lambda \mathbf{Q} | \mathbf{Q}\mathbf{Q}^T = \mathbf{I}, 0 \neq \lambda \in R\}$$

It is clear that the group of orthogonal transformations is contained in $\mathcal{G}_{zs}$. If $\mathcal{G}_{zs} = \mathcal{G}_\mathbf{O}$ then it is possible to calibrate cameras uniquely up to orthogonal transformation.

Theorem 1 *Let $\mathcal{G}_{zs}$ denote the class of transformations that preserve the zero-skew camera condition and $\mathcal{G}_\mathbf{O}$ the group of orthogonal transformations. Then*

$$\mathcal{G}_{zs} = \mathcal{G}_\mathbf{O}.$$

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

$$\mathbf{KRUQV} = \mathbf{KR}''\mathbf{QV} = \mathbf{K}'\mathbf{R}'''\mathbf{V} = \mathbf{K}'\mathbf{R}'$$

where $\mathbf{R}''$ and $\mathbf{R}'''$ denote orthogonal matrices. Now, using singular value decomposition of $\mathbf{Q}$ we may write

$$\mathbf{D} = \mathbf{UQV} = \begin{bmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{bmatrix}.$$

Suppose that the rotation matrix $\mathbf{R}$ is given by θ degree rotation about x-axis and by ϕ degrees about y-axis (note that the rotation $\mathbf{R}$ is arbitrary)

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

we have

$$\mathbf{RD} = \begin{bmatrix} d_1\cos\phi & d_2\sin\theta\sin\phi & -d_3\cos\theta\sin\phi \\ 0 & d_2\cos\phi & d_3\sin\phi \\ d_1\sin\phi & -d_2\sin\theta\cos\phi & d_3\cos\theta\cos\phi \end{bmatrix}.$$

Now, according to Lemma 1, $\mathbf{RD} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3]^T$ is a zero-skew calibration matrix if and only if $(\mathbf{r}_1 \times \mathbf{r}_3) \cdot (\mathbf{r}_2 \times \mathbf{r}_3) = 0$. After some calculation we have

$$\mathbf{r}_1 \times \mathbf{r}_3 = [0, -d_1 d_3 \cos\theta, -d_1 d_2 \sin\theta]^T$$
$$\mathbf{r}_2 \times \mathbf{r}_3 = [d_2 d_3 \cos\phi, d_1 d_3 \sin\theta\sin\phi, -d_1 d_2 \cos\theta\sin\phi]^T$$

and

$$\begin{aligned}(\mathbf{r}_1 \times \mathbf{r}_3) \cdot (\mathbf{r}_2 \times \mathbf{r}_3) &= -d_1^2 d_3^2 \cos\theta\sin\theta\sin\phi + d_1^2 d_2^2 \cos\theta\sin\theta\sin\phi \\ &= d_1^2 \cos\theta\sin\theta\sin\phi(-d_3^2 + d_2^2) \\ &= 0.\end{aligned}$$

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

3. Estimation Method

From the equation (1), we have $\mathbf{H}_k\mathbf{K}_0\mathbf{K}_0^T\mathbf{H}_k^T = \mathbf{K}_k\mathbf{K}_k^T$, and if we know the principal points (x_k, y_k), other calibration parameters (α_k, β_k) can be computed using a linear equations. In other words, the scale parameters (α_k, β_k) may be parameterized by the principal points, and given principal points the scale parameters are linearly computed. Now we define a nonlinear error function to find the optimal calibration parameters including the principal points. Using the relationship

$$\mathbf{R}_k = \frac{\mathbf{K}_k^{-1}\mathbf{H}_k\mathbf{K}_0}{det(\mathbf{K}_k^{-1}\mathbf{H}_k\mathbf{K}_0)^{\frac{1}{3}}} \quad (3)$$

we minimize the following error function

$$E = \sum_{k=1}^{N} \left(\left\| \mathbf{R}_k\mathbf{R}_k^T - \mathbf{I}_{3\times 3} \right\|_F^2 + \left\| \mathbf{R}_k^T\mathbf{R}_k - \mathbf{I}_{3\times 3} \right\|_F^2 \right) \quad (4)$$

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 $x_k \leftarrow \bar{x}_k$ and $y_k \leftarrow \bar{y}_k$ for $k = 0, ..., N$
2. Compute α_k and β_k, for $k = 0, ..., N$
3. Compute the error E defined by equation (4).
4. if E is smaller than the previous one, record the calibration parameters
5. repeat 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

Noise level (σ in pixel)	f_0	f_1	u	v	r_x	r_y	r_z
0	1000	1100	330	230	10°	10°	0°
0.5 (mean)	1001.4	1101.7	328.6	228.3	9.96	9.99	-0.01
(σ)	15.0	16.9	9.0	9.5	0.22	0.19	0.07
0.7 (mean)	997.2	1097.0	331.7	231.6	10.05	10.01	-0.02
(σ)	21.9	23.8	13.4	13.0	0.28	0.25	0.08
1.0 (mean)	1005.1	1106.5	330.3	229.0	9.95	10.06	-0.01
(σ)	44.7	49.5	19.3	22.8	0.43	0.40	0.11

표 1: Computation results after 100 runs at each noise level. Rotation angles are 10°, 10° and 0° about x-axis, y-axis and z-axis, respectively. About one hundred matching points are used in the computation of the homography.

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 σ. 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 $(-3.78°, -10.27°, -0.57°)$.

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.

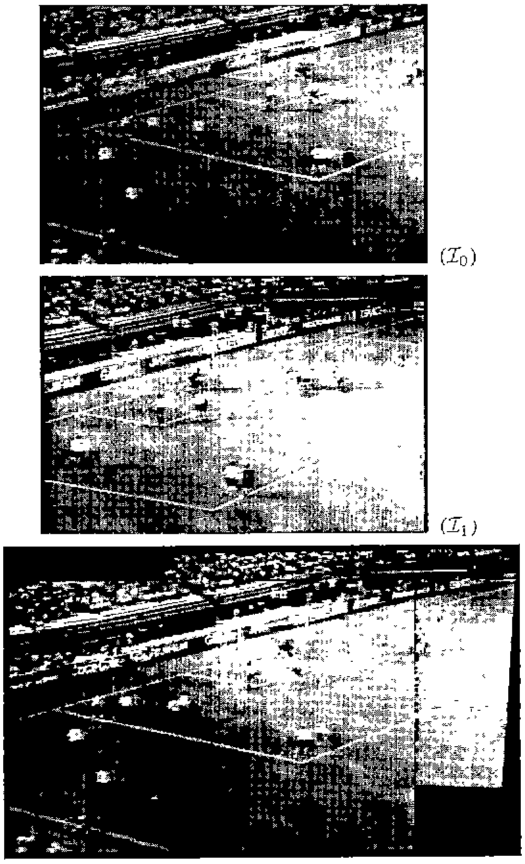

그림 1 Two images sampled from a real video of a soccer game. Using the homography computed, the mosaic image is obtained. The result is basically needed in estimating 3D locus of the ball or in virtual view synthesis.

참고문헌

[1] Richard I. Hartley. Self-calibration of stationary cameras. *International Journal of Computer Vision*, 22(1), 1997.

[2] Anders Heyden and Kalle Aström. Euclidean reconstruction from constant intrinsic parameters. In *Proc. ICPR'96*, pages 339–343, 1996.

[3] Anders Heyden and Kalle Astrom. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Proc. CVPR'97*, 1997.

[4] Anders Heyden and Kalle Astrom. Minimal conditions on intrinsic parameters for euclidean reconstruction. In *Proc. Asian Conference on Computer Vision, Hong Kong*, 1998.

[5] Stephen J. Maybank and Olivier D. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.

[6] Marc Pollefeys and Luc Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proc. Int. Conf. on Computer Vision*, 1998.

[7] Richard Szeliski and Sing Bing Kang. Direct methods for visual scene reconstruction. In *IEEE Computer Society Workshop: Representation of Visual Scenes*, 1995.

[8] B. Triggs. Autocalibration and the absolute quadric. In *Proc. CVPR'97*, 1997.