A variational approach to camera motion smoothing

Luis Alvarez* Luis Gómez† Pedro Hénriquez* and Luis Mazorra*

In this paper we deal with camera motion smoothing. We focus our attention on the case of video cameras mounted on a tripod. In such case, for each time t of the video sequence, camera motion configuration is provided by 3 parameters : (1) $P(t)$ (PAN) which represents the tripod vertical axis rotation, (2) $T(t)$ (TILT) which represents the tripod horizontal axis rotation and (3) $Z(t)$ (CAMERA ZOOM) the camera lens zoom setting. From these 3 parameter functions we can easily deduce (using tripod information), for each time t, the camera calibration, that is, the camera intrinsic and extrinsic parameters which determine the position of the camera in the $3D$ space and the way $3D$ objects are projected in the camera projection plane (the CCD in the case of digital cameras).

Human being visual system is very sensitive to motion, and small perturbations over time of $P(t)$, $T(t)$, and $Z(t)$ values produce small oscillations in the camera motion disturbing the observer. To remove such perturbations is a critical issue in applications like the inclusion of artificial graphic objects in real video sequence scenarios. To illustrate this phenomenon, in http://www.ctim.es/demo102 we show a real video sequence where some graphic objects have been included. In this video we illustrate the main practical problem we deal with, that is, standard calibration techniques which do not take into account the expected time regularity of $(P(t), T(t), Z(t))$ can introduce a significant noise in the camera motion estimation over time.

In this paper we propose to smooth $(P(t), T(t), Z(t))$ functions by minimizing the following energy :

$$E(P(t), T(t), Z(t)) = \int_{t_0}^{t_1} \left(w_P P'(t)^2 + w_T T'(t)^2 + w_Z Z'(t)^2 + F(P(t), T(t), Z(t), t) \right) dt, \qquad (1)$$

where $[t_0, t_1]$ is the time interval, $w_P, w_T, w_Z \geq 0$ are weights to balance the different components of the energy and $F(x_1, x_2, x_3, t) \geq 0$ is an standard calibration function which forces, for each time t, that the projection of $3D$ points be close to primitives detected in the image. In fact, when no time regularization is used, $(P(t), T(t), Z(t))$ is usually estimated by minimizing $F(P(t), T(t), Z(t), t)$ independently for each time t. So, by using the proposed variational model (1), we introduce a time regularity condition in the video camera calibration procedure. In http://www.ctim.es/demo102 we illustrate the results we obtain by applying the proposed method. We include the smoothed video sequence and some graphics to compare original and smoothed $(P(t), T(t), Z(t))$ functions

The Euler-Lagrange equations associated to energy (1) yields to the following nonlinear system of differential equations:

$$\begin{cases} -w_P P''(t) + \frac{\partial F}{\partial x_1}(P(t), T(t), Z(t), t) = 0 \quad in \quad (t_0, t_1) \\ -w_T T''(t) + \frac{\partial F}{\partial x_2}(P(t), T(t), Z(t), t) = 0 \quad in \quad (t_0, t_1) \\ -w_Z Z''(t) + \frac{\partial F}{\partial x_3}(P(t), T(t), Z(t), t) = 0 \quad in \quad (t_0, t_1), \end{cases}$$

with adequate boundary conditions. In this paper we study the minimization problem (1) and we present some numerical experiments to illustrate the smoothing performance of the proposed method.

*CTIM. Departamento de Informática y Sistemas. Univ. Las Palmas de G.C., Campus de Tafira, 35017 Las Palmas de G.C. e–mail lalvarez@ctim.es phenriquez@ctim.es lmazorra@ctim.es

†CTIM. Departamento de Ingeniería Electrónica y Automática. Univ. Las Palmas de G.C., Campus de Tafira, 35017 Las Palmas de G.C. e–mail lgomez@ctim.es