

선형 중첩된 광류 모델을 이용한

카메라 움직임 파라미터 추출

박상철[○], 한성식, 이성환
(주)버추얼미디어[○], 고려대학교 인공지능연구센터
scpark@virtualmedia.co.kr / {sshan, swlee}@image.korea.ac.kr

Extraction of Camera Motion Parameters with Linear Composition of Optical Flow Models

Sang-Cheol Park[○], Seong-Sik Han, Seong-Whan Lee
Virtual Media, Inc.[○], Center for Artificial Vision Research, Korea University

요 약

본 논문에서 제안하는 카메라 움직임 파라미터를 추출하는 새로운 방법은 먼저 광류(optical flow) 관계 수식을 이용하여, 카메라의 다양한 움직임에 의해 생성되는 광류 모델들을 생성한다. 입력 비디오 데이터의 인접 영상으로부터 산출된 광류를 생성된 광류 모델들을 이용하여 선형 분해한다. 분해 과정을 통해 각 모델에 대한 가중치를 산출할 수 있으며, 산출된 가중치는 각 모델의 생성에 사용된 파라미터에 적용하여 이들의 선형 조합을 통해 입력에 대한 카메라 움직임 파라미터를 추출할 수 있다. 제안된 방법의 성능을 평가하기 위하여 수치 데이터와 비디오 데이터에 대하여 실험한 결과, 제안된 카메라 움직임 파라미터 추출 방법이 적은 계산 비용으로 정확하게 카메라 움직임 파라미터를 추출할 수 있음을 확인할 수 있었다. 또한 입력 데이터에 노이즈가 포함된 경우에도 파라미터 추출 성능이 우수함을 알 수 있었다.

1. 서론

통신 기술 및 하드웨어의 발달로 문자나 숫자 등과 같이 단일한 형태의 데이터로부터 동영상과 같이 다양한 형태의 데이터인 멀티미디어로의 관심이 집중되고 있다. 특히, 급증하고 있는 비디오 데이터를 효율적으로 분석하여 보다 체계적이고 구조적인 형태로 재구성하고 이를 관리해야 할 필요성이 절실히 요구되고 있다.

일반적으로 동영상 데이터에는 보다 효과적인 정보 전달을 위해 다양한 카메라 조작 기술이 사용되어 진다. 관심 객체의 이동 방향에 따라 카메라의 조작이 이루어지며, 카메라 움직임 파라미터 추출이란 동영상 데이터로부터 카메라 조작 정보를 추출하는 것을 의미한다.

카메라 움직임 분석 방법은 분석 환경에 따라서 주파수 영역과 공간 영역에서의 카메라 움직임 분석으로 나뉘어 진다. 주파수 영역에서의 카메라 움직임 분석은 연구 초기에 많이 사용된 방법으로 Maeda 등[1]은 일정 크기의 연속된 영상으로부터 시공간 영상(Spatio-temporal image)을 생성하여 이를 DFT(Discrete Fourier Transform)을 통해 나타나는 정보를 분석하여 카메라 움직임 정보를 추출하였고, Joly[2]는 시공간 영상으로부터 선 성분을 추출하여 카메라의 움직임을 분석하였다. 공간 영역에서의 분석 방법의 경우, Park 등[3]은 인접 영상간 대응관계 정보를 원근 사상 모델과 회전, 확대/축소 모델에 기반한 변형 모델에 도입하여 카메라의 움직임 파라미터를 추출하였다.

본 논문에서는 동영상으로부터의 카메라 움직임 파라미터 추출 방법을 제안한다. 제안된 방법에서는 광류 관계 수식을 이용하여 노이즈가 포함되지 않은 이상적인 광류

모델들을 생성한 후, 입력으로부터 계산된 광류를 생성된 광류 모델 내에서 분석하여 입력에 대한 카메라 움직임 파라미터를 추출하므로, 입력 영상으로부터 광류 정보를 계산하는 과정에서의 일부 오차와 입력 영상의 일부 영역이 노이즈를 포함한 경우에도 비교적 정확한 카메라 움직임 파라미터를 추출할 수 있다.

2. 카메라 시스템

2.1 카메라 모델

카메라는 3차원의 실세계의 점 좌표를 2차원 영상 평면상의 좌표로 투영하여 나타낸다. 그림 2는 카메라 중심 좌표를 갖는 일반적인 카메라 모델을 나타낸다.

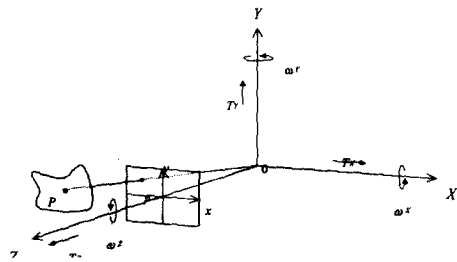


그림 1. 카메라 모델

P 는 3차원 실세계 임의의 점의 위치를 나타내며 x, y 평면은 카메라의 움직임에 의해 3차원상의 임의의 점 P 에 의해 생성되는 광류가 나타나는 영상 평면이다. 카메라는 실세계 점의 좌표 $P=(X, Y, Z)$ 를 2차원 평면의 좌

표 $p=(x, y)$ 로 투영한다. 카메라의 중심으로부터 영상평면까지의 거리는 카메라의 초점 거리 f 가 되며, 이때 3차원에서 2차원에서의 사상 관계는 식 (1)과 같이 정의된다.

$$x = f \cdot \frac{X}{Z}, \quad y = f \cdot \frac{Y}{Z} \quad (1)$$

2.2 카메라 움직임에 의한 광류

카메라의 시계가 $20^\circ \times 30^\circ$ 에 비해 충분히 작은 상황에서 카메라의 움직임에 의해 생성되는 광류는 다음과 같은 특성을 갖는다. 카메라의 수평선회(Panning)에 의해 생성되는 광류는 가로 방향성의 벡터로 이루어지며, 상하선회(Tilting)에 의해 생성되는 광류는 세로 방향성의 벡터로 구성되어진다.

카메라의 회전($\Omega_x, \Omega_y, \Omega_z$), 이동(T_x, T_y, T_z), 확대/축소(r_{zoom}) 모션에 따라 3차원 상의 점들의 움직임은 영상 평면에서 광류로 관계되어 나타내어지는데, 이때 생성되는 광류는 x 성분 $u(x, y)$ 과 y 성분 $v(x, y)$ 을 갖는 벡터로 식 (3), (4)와 같이 정의된다[4].

$$u(x, y) = -\frac{f}{Z}T_x + \frac{xy}{f}\Omega_x - f\left(1 + \frac{x^2}{f^2}\right)\Omega_z + f\left[\tan^{-1}\left(\frac{x}{f}\right)\right]\left(1 + \frac{x^2}{f^2}\right)r_{zoom} \quad (2)$$

$$v(x, y) = -\frac{f}{Z}T_y - \frac{xy}{f}\Omega_x + f\left(1 + \frac{y^2}{f^2}\right)\Omega_z + f\left[\tan^{-1}\left(\frac{y}{f}\right)\right]\left(1 + \frac{y^2}{f^2}\right)r_{zoom} \quad (3)$$

3. 카메라 움직임 파라미터 추출 알고리즘

3.1 카메라 움직임에 의한 광류 모델

본 논문에서는 카메라 움직임 파라미터들의 조합으로 생성된 광류 모델을 이용하여 입력 비디오에 대한 카메라 움직임 파라미터를 추출한다. 위에서 제시한 카메라의 시계가 충분히 작은 상황을 고려하여 u, v 의 카메라 회전 성분 u^{rot}, v^{rot} 에 대해 다음의 식 (4), (5)로 정의할 수 있다.

$$v(x, y) = -\frac{f}{Z}T_y - \frac{xy}{f}\Omega_y + f\left(1 + \frac{y^2}{f^2}\right)\Omega_x - x\Omega_z \quad (4)$$

$$\cong -f\Omega_y + y\Omega_z + xr_{zoom}$$

$$u(x, y) = -\frac{f}{Z}T_x + \frac{xy}{f}\Omega_x - f\left(1 + \frac{x^2}{f^2}\right)\Omega_y + y\Omega_z \quad (5)$$

$$\cong f\Omega_x - x\Omega_z + yr_{zoom}$$

3.2 입력 광류 및 광류 모델 데이터 생성

파라미터 추출을 위해 입력 비디오 데이터로부터 산출된 광류 정보가 된다. 정확한 광류 정보를 제공하기 위해 입력 비디오 데이터로부터 추출된 에지 정보를 이용하여 특징점을 추출하고, 그 특징점에서 산출된 광류 정보의 이용하여 광류 데이터를 생성한다.

또한, 입력 영상과 같은 크기의 참조 모델에 대해 산출된 광류 데이터의 모든 위치를 이용하여 광류 모델을 생성하고, 이 광류 모델로부터 카메라 파라미터 추출을 위해 사용될 정보를 선별하여 사용한다.

3.3 선형 중첩된 광류 모델을 이용한 파라미터 추출

독립적인 광류 모델의 정보를 가지고 광류 모델의 평균과 각 광류 모델과의 차이 $\tilde{T} = T - \bar{T}$ 의 공분산 C_T 를 구하고, 주성분 분석(Principal Component Analysis)을 사용하여 m 개의 광류 모델에 대한 고유 벡터 t_i 를 기저로 하는 2차원 광류 모형을 얻는다[5][6].

$$T = \bar{T} + \sum_{i=1}^{m-1} \alpha_i t_i \quad (6)$$

입력 영상 생성에 있어서, 단지 영상의 특징점에 대한 광류 정보만이 주어지기 때문에 요구되는 파라미터의 근사치만을 추정할 수 있다. 따라서 최적해를 찾기 위해 먼저 정규화된 계수들의 합을 에너지 함수로 정의하고, 주어진 특징점에서의 광류 정보가 완전히 복원 가능해야만 한다는 조건을 설정한다. 에너지 함수 $E(\alpha)$ 는 평균 광류로부터 변형 정도를 나타낸다. 본 논문에서 해결하고자 하는 문제는 조건을 만족하는 계수들 가운데 $E(\alpha)$ 를 최소화하는 α 를 찾는 것이다[5].

$$\alpha^* = \arg \min_{\alpha} E(\alpha) \quad (7)$$

$$E(\alpha) = \sum_{i=1}^{m-1} \left(\frac{\alpha_i}{\sigma_i}\right)^2 \quad (8)$$

$$\tilde{T}(x_j) = \sum_{i=1}^{m-1} \alpha_i t_i(x_j) \quad (j = 1, \dots, n) \quad (9)$$

이때 x_1, \dots, x_n 는 선택된 특징점들이며 기저의 수에 비해 적은 수의 특징점을 선택함으로 n 은 $(m-1)$ 보다 작다.

식 (7)~(9)에 의해 이 문제는 최소 제곱법을 이용하여 단순화할 수 있다.

최종적으로 전체 영상에 대한 정보를 얻기 위해서는 다음과 같은 식으로 나타내게 된다.

$$T = \bar{T} + \sum_{i=1}^{m-1} \alpha_i^* \sigma_i t_i \quad (10)$$

4. 실험 결과 및 분석

4.1 실험 환경

비디오 데이터는 SONY Digital Cam(DCR-PC3)으로 촬영하여 Optibase MPEG Fusion System MPEG-2 Encoder를 이용하여 초당 15프레임 이상의 샘플링 비율로 생성하였다. 파일 형식은 MPEG2이며, 각 프레임은 704×480 해상도의 RGB 컬러 영상으로 구성되어 있다.

4.2 실험 결과

4.2.1 노이즈가 포함되지 않은 데이터의 실험 결과

생성된 광류 데이터를 제안된 알고리즘의 입력으로 하여, 광류 데이터 생성시에 사용된 카메라 움직임 파라미터와 알고리즘을 통해 추출된 파라미터 값의 비교를 통해 알고리즘의 성능을 평가한다.

표 1은 광류 관계 수식을 통해 광류 데이터 생성에 사용된 파라미터와 제안된 알고리즘을 통해 추출된 카메라 움직임 파라미터 값을 나타낸다.

구분 \ 파라미터	T_x	Ω_y	T_y	Ω_x	Ω_z	r_{zoom}
입력 Optical Flow 생성 파라미터	-0.2794	0.1947	0.0970	-0.4765	-0.9014	0.1421
예측 파라미터	-0.1010	0.0163	0.0902	-0.4832	-0.9014	0.1421
입력 Optical Flow 생성 파라미터	-0.5457	0.4064	0.0325	-0.0836	0.1650	0.0184
예측 파라미터	-0.1255	-0.0138	0.0578	-0.0583	0.1650	0.0184

T_x : 좌우 이동 T_y : 상하 이동 Ω_z : 광축 회전
 Ω_y : 상하 선회 Ω_x : 수평 선회 r_{zoom} : 확대/축소

표 1. 노이즈가 포함되지 않은 광류 데이터의 성능 실험

4.2.2 노이즈가 포함된 데이터의 실험 결과

생성된 광류 데이터에 노이즈를 점차적으로 포함시켜 제안된 알고리즘의 노이즈에 대한 강건성을 실험한다.

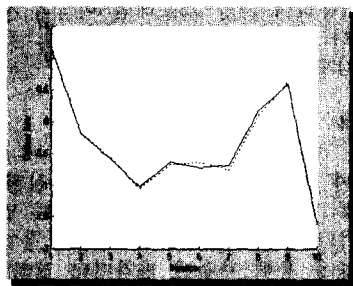
구분 \ 파라미터	T_x	Ω_y	T_y	Ω_x	Ω_z	r_{zoom}
입력 Optical Flow 생성 파라미터	-0.3409	-0.2042	0.9073	0.1467	-0.7061	-0.4438
예측 파라미터	-0.3078	-0.2722	0.4265	-0.3065	-0.7054	-0.4441

표 2. 노이즈 포함비율 8%에 대한 성능 실험

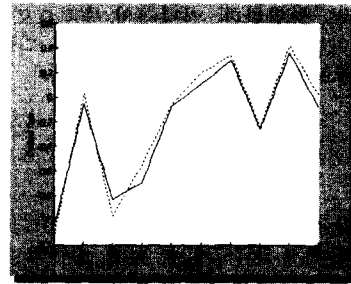
구분 \ 파라미터	T_x	Ω_y	T_y	Ω_x	Ω_z	r_{zoom}
입력 Optical Flow 생성 파라미터	0.6125	-0.8998	-0.1990	0.9564	-0.9337	0.9800
예측 파라미터	-0.0853	-0.4114	-0.5927	0.5051	-0.9289	0.9770

표 3. 노이즈 포함비율 10%에 대한 성능 실험

그림 3은 광류의 파라미터 값(Optical Flow: y축)과 수행회수(iteration: x축)로 도식화한 그래프이다. 실제 광류 데이터에 의한 파라미터와 제안 알고리즘에 의해 산출된 광류 파라미터의 비교 성능 평가 역시 노이즈의 증가에도 큰 차이가 나지 않는 노이즈에 대한 강건성을 보이고 있다.



(a) 노이즈 포함 비율 8%



(b) 노이즈 포함 비율 10%

— 실제 값
 예측 값

그림 3. 노이즈 포함비율에 따른 비교 성능 평가

5. 결론 및 향후 연구 방향

본 논문에서는 동영상으로부터 카메라 움직임 파라미터를 추출하는 방법을 제안하고, 그 성능을 평가해 보았다. 카메라의 움직임에 의해 생성되는 광류와의 관계식을 통해 광류 데이터를 생성하고 이를 이용하여 실험한 결과 노이즈가 포함되지 않은 데이터에 대해서는 실제의 움직임 파라미터와 오차 없이 추출해내는 성능을 보였으며, 노이즈를 점차적으로 추가하여도 실제 움직임 파라미터에 가까운 예측 성능을 보임을 확인할 수 있었다.

카메라 움직임 파라미터 추출 기술은 나아가 동영상 데이터 상의 객체 분할, 객체 추적, 비디오 인덱싱, 비디오 압축, 증강 현실과 같은 응용 분야의 핵심 요소 기술로 적용되어 수행되어야 할 것이다.

참고문헌

- [1] J. Maeda, "Method for extracting camera operations to describe sub-scenes in video sequences," *Proc. of SPIE*, Vol. 2187, pp. 56-67, 1994.
- [2] P. Joly and H.-K. Kim, "Efficient automatic analysis of camera work and microsegmentation of video using spatiotemporal images," *Signal Processing*, Vol. 8, No. 4, pp. 295-307, 1996.
- [3] J. Park, N. Yagi, K. Enami and Kiyoharu, "Estimation of camera parameters from image sequence for model based video coding," *IEEE Trans. Circuits Systems Video Tech*, Vol. 4, No. 3, pp. 288-296, 1994.
- [4] M. V. Srinivasan, S. Venkatesh and R. Hosie, "Qualitative estimation of camera motion parameters from video sequences," *Pattern Recognition*, Vol. 30, pp. 593-606, 1997.
- [5] B.-W. Hwang, V. Blanz, T. Vetter, H. -H. Song and S. -W. Lee, "Face Reconstruction using a Small Set of Feature Points," *Proc. of 1st IEEE International Workshop on Biologically Motivated Computer Vision*, pp. 308-315, Seoul, Korea, May 2000.
- [6] T. Vetter and T. Poggio, "Linear object classes and image synthesis from a single example image," *IEEE Trans. Image Processing*, Vol. 19, No. 7, pp. 733-742, 1997.