

Periodic Motion의 Invariant Detection을 위한 Affine Model 적용

최우진, 정진현
광운대학교 제어계측공학과

Invariant Detection of Periodic Motion using Affine Model

Woo Jin Choi, Chin Hyun Chung
Department of Control and Instrumentation Eng. Kwangwoon University

Abstract - A limitation is assumed that In this paper, a generalized method is proposed to extract a period of a motion of an object. To detect a periodic motion, we put restrictions on a stationary camera and on a motion of an object. We can derive the necessary and sufficient condition that an image sequence consists of the projection of the periodic motion by the affine transformation that is a reasonably good approach to the perspective projection. The difficulty of detecting its periodic motion is to select its have period in sequence and to define its width.

하고 있다고 하자. 직관적으로 jogging motion이 Eq. (1)을 완벽히 만족하지 않아도 반복, 순환운동처럼 보일 것이다. 즉, 팔과 다리를 일정한 각도와 크기로 움직이고, 머리카나 다리 등의 point들이 지면을 기준으로 했을 때 정형 파와 같은 곡선을 그리는 것을 볼 수 있다. 그 사람은 신체의 일정한 움직임과 임의의 경로의 두 가지 component로 나눌 수 있다.

1. 서 론

지구상의 모든 운동을 한번 생각해보자. 우리가 보는 대부분의 움직임이 실제로 주기적인 성분을 가지고 있다. 즉, 대부분의 locomotory motion(달리기, 경보, 수영 등등)을 살펴보면 움직임들이 repeat and translatory component들로 구성되어있다는 것을 알 수 있다. periodic motion을 얻기 위해서는 주기적인 성분을 뽑아내야 하는데 문제는 camera가 움직임에 따라 view-point가 변하기 때문에 periodic motion의 projection이 변한다는 것이다. 많은 움직임들이 periodic component를 가지고 있지만 그것들의 size와 scale에서는 조금씩 다르고 Image sequence의 주기는 각 cycle마다 다를 수 있다. 즉, cycle마다 base period의 scaling이나 shifting이 포함된 base period와 유사한 다른 주기를 가지고 있다. 이 연구에서는 object의 scale, orientation, position과 view-point가 변한다는 가정을 둔다. 이 가정은 affine transform으로 modeling할 수 있다. 그래서 affine invariance의 결과(1)를 이용할 수 있다.

2.1.1 Affine transformation.

Eq. (1)은 주기성에 영향을 주지 않는 transformation을 이용해서 부분적으로 주기적인 motion을 describe할 수 있게 한다. 본 연구에서는 affine transformation을 이용한다.

$$M(t) = A(t) \cdot C(t) \quad (2-2)$$

C는 purely-periodic motion이고 A가 time-varying affine transformation일 때 위와 같은 식이 만족한다면 motion M은 affinely-periodic이라 한다. affine transformation은 rigid transformation의 modeling뿐만 아니라 계산하기 편리한 property를 가지고 있다. projection 과정이 invertible하지 않기 때문에 모든 3D periodic motion에 대해 detection하는 것을 바랄 수는 없다. 그러나 projection motion은 3D periodic motion이 주기적으로 나타나지 않아도 주기성을 보여주고 그 반대도 가능하다. 문제는 image sequence가 time상에서 sampling되는 것이다. 최선의 방법은 projected sampled motion이 3D periodic motion에서 나와야 하는 것이다. 이것을 명백히 하기 위해서 frame이란 용어를 사용하겠다고 motion C(t)의 projection을 나타내기 위해서 notation F,를 사용하겠다. Frame이 time-sampling된 affinely-periodic motion의 orthographic projection이라면 이 frame sequence를 apparently-periodic이라고 한다.

2. 본 론

2.1 Periodic Motion

Motion M이 주기 p를 가지고 있을 때 purely-periodic이라 하고 다음과 같이 나타낸다.

$$M(t+p) = M(t) \quad (2-1)$$

M(t)의 ith column은 time t에서 ith point의 3D position이고 p는 주기를 나타내고 양수이다.

purely-periodic motion의 예를 들자면 선풍기의 날개, helicopter의 propeller, J의 회전운동 등이 있다. camera가 정지해있을 때 R³에서의 purely-periodic motion은 R²로 projection하면 R²에서 purely-periodic이다. periodic motion의 좀더 발전된 case는 부분적으로 periodic이거나 camera가 움직이는 경우이다. 예를 들어 어떤 사람이 운동장에서 임의의 경로로 jogging을

2.2 Affine Matching

projected, affinely-periodic motion의 다른 cycle에 상응하는 frame의 관찰 결과에 입각한 접근은 affine transformation을 기반으로 한다. matching에 대한 개념은 motion이 apparently-periodic이란 필요충분조건이다. 이 section에 periodic motion에 관한 matching에 대한 것과 matching이 무엇을 의미하는지 정의한다.

2.2.1 3D Match Criteria

m 3D shape가 모든 single shape의 affine transformation이라면 3D shape는 matching되었다고 말한다. motion C가 time t에 대해 shape S ∈ R³ × Rⁿ와 affine transformation {A_k} ⊂ R³ × R³가 존재하면 C는 affinely-periodic(주기 p)이라 한다. 이 경우에 모든 정수 k에 대해 C(t+kp) = A_k · C이다. periodic motion의 다른 cycle의 상응하는 frame은 같은 shape의 affine transformation이다. [1]에서

보여준 결과대로 affinely corresponding frame의 저장된 연쇄에 의해 만들어진 matrix는 거의 rank가 3이다.

$$M_i^p = \begin{bmatrix} F_i \\ F_{i+p} \\ \vdots \\ F_{i+(k-1)p} \end{bmatrix} \quad (2-3)$$

measurement matrix는 각 원소의 행에서 각 행의 중심을 뺀으로써 나타난 효과를 제거시키기 위해서 표시한다. M_i^p 의 registered version은 \widehat{M}_i^p 이다. 3D matching criterion과 [1]에 의해 \widehat{M}_i^p 은 다음과 같이 표현한다.

$$\widehat{M}_i^p = RS \quad (2-4)$$

R은 camera rotation으로 R의 행은 stream을 통해 나오는 camera 수평, 수직 reference 축의 방향으로 크기는 $2k \times 3$ 이고, S는 shape matrix로 S의 열은 object의 중심을 기준으로 하여 feature point의 좌표로 행렬의 크기는 $3 \times n$ 이다. noise가 없을 때 registered measurement matrix \widehat{M}_i^p 의 rank는 3이거나 그보다 작다. 결과적으로 \widehat{M}_i^p 는 SVD(singular value decomposition)로 rank가 3이거나 그보다 작은 행렬로 분해할 수 있다.

$$\widehat{M}_i^p = O_1 \Sigma O_2 \quad (2-5)$$

이 식은 \widehat{M}_i^p 의 SVD로서 O_1 은 $2k \times 3$ 이고, O_2 는 $3 \times n$ matrix이다.

2.2.2 Projected Match Criterion

Frame sequence F_1, \dots, F_m 는 \widehat{M}_i^p 가 $i=1, \dots, m$ 에서 rank가 3일 때 apparently-periodic이다. frame의 집합 $\{F_{i+jp}\}_{j=0}^{k-1}$ 는 \widehat{M}_i^p 의 rank가 3이거나 작을 때 matching되었다고 한다. 즉, frame 집합이 같은 object의 affine transformation의 projection으로 표현될 때 matching되었다고 한다.

2.2.3 Approximation Matching

frame 집합이 matching이 안될 때 matching을 시키기 위해서 frame과의 어떤 양에 의한 상대적인 거리로 표현할 수 있다. frame 집합의 거리는 다음과 같이 나타낸다.

$$dist_A(\{F_{i+jp}\}_{j=0}^{k-1}) = \min(\|E\|_{rms}) \quad (2-6)$$

단, $rank(\widehat{M}_i^p + E) \leq 3$ 이고 $\|E\|_{rms}$ 는 matrix E의 root-mean-squared norm이다.

즉, $\|E\|_{rms} = \sqrt{\frac{1}{2kn} \sum_{i,j} E_{ij}^2}$. distance에 대한 정의는 계산상으로 유용하지 않지만 \widehat{M}_i^p 의 singular value term으로 $dist_A(\{F_{i+jp}\}_{j=0}^{k-1})$ 를 다시 나타내는 간단한 방법이 있다[6]. 즉,

$$dist_A(\{F_{i+jp}\}_{j=0}^{k-1}) = \sqrt{\frac{1}{2kn} \sum_{i=4} \sigma_i^2} \quad (2-7)$$

여기서 σ_i 는 \widehat{M}_i^p 의 i th singular value이다.

$dist_A(\{F_{i+jp}\}_{j=0}^{k-1})$ 는 완벽한 matching을 위해 frame 집합에서 각 feature의 2D 위치가 교란되지 않도록 필요한 평균양(픽셀)을 준다. frame X, Y의 distance는 $dist_A(\{X, Y\})$ 로 나타내고 $dist_A(X, Y)$ 로

줄일 수 있다. 위 식을 간단히 distance라고 하고 다음은 여기에 대한 다섯 가지 정의이다.

- ① frame의 집합, $\{X_i\}$ 이 정확히 matching이 되면 $dist_A(\{\{X_i\}^m\}) = 0$ 이다.
- ② distance는 feature measurement perturbation이란 noise에 대해 비교적 영향이 적다.
- ③ distance는 image coordinate에서 정의되고 measurement error에 대해 직접적으로 관계된다.
- ④ distance는 feature가 5보다 적을 경우 항상 영이다.
- ⑤ distance의 계산은 수학 연산자, $O(n)$ 가 필요하다. m feature에 대해 계산량은 $O(nm^2)$ 과 $O(mn^2)$ 의 최소이다.

2.3 Algorithm

모든 i 값에 대해 \widehat{M}_i^p 의 p 값이 rank 3이거나 그보다 작으면 motion은 p 주기 apparently-periodic이다. 위의 계산치는 다음과 같은 confidence function으로 나타낸다.

$$confidence_f = \max(0, 1 - \frac{f(t)}{mean_f}) \quad (2-8)$$

confidence function은 $mean_f$ 의 영일 때 한 개의 singular value를 갖는다. 이 경우에 motion은 주기 1을 갖는다. motion이 주기적이 그것의 주기를 판별하기 위한 algorithm은 다음과 같다.

- ㉠ $i < j$ 의 모든 값에 대해 $dist_A(F_i, F_j)$ 를 계산한다.
- ㉡ $p=i-j$ 의 각 값에 대해 ㉠에서 구한 결과를 평균한다. 즉, $M(p) = \frac{mean}{j=i-k^*p} \{ dist_A(F_i, F_j) \}$ 를 계산한다.
- ㉢ 계산 값이 허용할 수 있는 주기인지를 판별하기 위해서 $confidence_M$ 을 계산한다.

위의 algorithm의 결과는 frame들로 얻어진 주기들의 값들의 confidence list로 나타내어진다. 실제로 가장 높은 confidence를 가지고 있는 주기 p_{max} 는 true period의 정수 곱에 상당한다. 위의 true period는 p_{max} 의 가장 작은 소수이다. ㉢의 과정은 Fourier Transformation을 사용할 수 있고 기본 주파수(주기를 줄일 수 있다. 그러나 이 경우에 이론적으로 식 (2-1)을 만족하지 못한다. 이것을 이용하기 위해서는 $M(k^*p) = 0$ 이라는 조건을 삽입해야한다.

3. 결 론

periodic motion의 주기를 찾기 위해 150 frame의 동영상을 입력 데이터로 사용하였다.

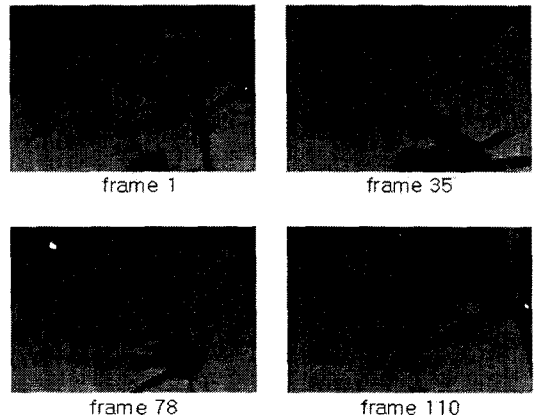


그림 A 야자수 영상 스트림(stream)

[참고 문헌]

150 frame의 영상에서 주기를 계산한 다음 그 결과의 신뢰도를 확인하기 위해서 식 (2-8)을 이용하여 다음과 같은 graph로 나타내었다.

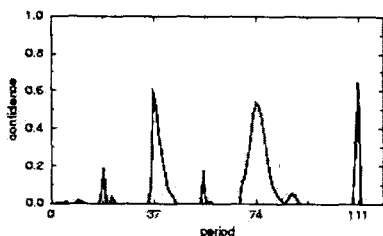


그림 B 입력 영상의 각 프레임에 대한 신뢰도

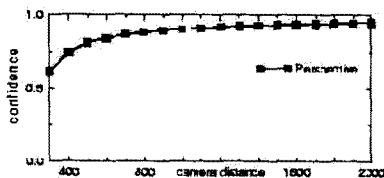


그림 C 카메라 거리에 따른 신뢰도

그림 B의 가로축은 period, 즉 frame이고 세로축은 Eq.(2-8)의 confidence값이다. 위의 그래프는 입력 영상에서 17 프레임을 가장 작은 주기로 볼 수 있으나 confidence값이 0.2에 미치지 못해 신뢰할 수 없다. 여기에서 우리는 confidence값이 0.5이상이어야 신뢰할 수 있다고 말할 수 있다. 따라서 가장 작은 주기로는 0.6의 신뢰도를 가지고 있는 37 프레임을 기본 주기로 잡을 수 있다. 위의 결과에서 가장 높은 신뢰도를 보여주는 것이 111 프레임이다. confidence값은 0.67이다. 따라서 프레임 111이 가장 신뢰할 수 있는 주기이다. 그림 C의 graph는 가로축은 카메라와 object와의 거리이고 세로축은 confidence값이다. 카메라와 object의 거리를 0에서 3000의 임의의 거리로 설정하고 신뢰도가 0.6이상에 대한 그래프이다.

- [1] C. Tomasi and T. Kanade, "Shape and Motion from image streams under orthography", Intl. Journal of Computer Vision, vol. 9, no. 2, pp. 137-154, 1992
- [2] S. M. Seitz and C. R. Dyer, "Affine invariant detection of periodic motion", Tech. Rep. 1225, University of Wisconsin, Madison, WI, 1994
- [3] S. M. Seitz and C. R. Dyer, "Complete Scene Structure from Four Point Correspondences" Intl. Conference of Computer Vision, pp. 330-337, 1995
- [4] S. M. Seitz and C. R. Dyer, "View-Invariant Analysis of Cyclic Motion", Intl. Journal of Computer Vision, vol 25, pp. 1-23, 1997
- [5] P.-S. Tsai and M. Shah, "Cyclic motion detection" Tech. Rep. 1225, CS-TR-93-08, University of Central Florida, Orlando, FL, 1993.
- [6] G. W. Stewart, "Introduction to Matrix Computations", New York, NY : Academic Press, 1973.
- [7] Jan J. Koenderink and Andrea J. van Doorn, "Affine structure from motion" Opt. Soc. Am. A, 8, 377-385, 1991.
- [8] Ramprasad Polana and Randal Nelson. "Detecting activities", In Proc. Computer Vision and Pattern Recognition, 2-7, 1993.
- [9] Daphna Weinshall, "Model-based invariants for 3D vision", Intl. Journal of Computer Vision, 10(1), 27-42, 1993.