

시점 불변 게이트 인식을 위한 호모그래피의 추정

나진영*, 강성숙**, 정승도*, 최병욱**

*한양대학교 전자통신전파공학과

**한양대학교 정보통신대학원

e-mail:youngly@ihanyang.ac.kr

Homography Estimation for View-invariant Gait Recognition

Jin-Young Na*, Sung-Suk Kang**, Seung-Do Jeong*, Byung-Uk Choi**

*Dept. of Div. of Electrical and Computer Engineering, Hanyang University

**Dept. of Information and Communications, Hanyang University

요 약

게이트는 사람의 걷는 방법 혹은 그 특성을 나타내는 용어로서, 최근 컴퓨터 비전 기술을 이용하여 개개인을 분별하기 위한 게이트 특징 정보를 추출하고자 하는 연구가 활발히 진행되고 있다. 그러나 영상을 기반으로 추출한 게이트 정보는 카메라의 시점에 종속적인 단점을 가지고 있다. 이러한 단점을 해결하기 위한 노력으로 3차원 정보를 획득하려는 연구가 진행되고 있으나, 이는 카메라와 사람간의 거리, 카메라 파라미터 등 부가적인 정보를 필요로 한다. 본 논문에서는 영상내의 정보만을 이용하여, 카메라 시점에 종속적인 게이트 인식의 단점을 해결할 수 있는 방안을 제안한다. 먼저 실루엣 영상으로부터 걷는 방향을 찾아내고, 간단한 연산을 통해 평면 호모그래피를 추정한다. 추정된 호모그래피를 이용하여 측면 시점의 영상으로 재구성하면, 시점 변화에 비종속적인 게이트 정보를 추출할 수 있다. 본 논문에서 제안한 방법을 평가하기 위하여 실루엣 영상의 폭과 높이 변화를 비교하였다. 실험을 통해 제안한 방법을 적용할 경우, 그렇지 않은 경우에 비하여 특징 변화가 적음을 확인하였고, 특히 보폭 등의 게이트 특징 정보가 일정한 값을 유지함을 볼 수 있었다.

1. 서론

게이트(Gait)는 사람의 걷는 방법을 지칭하는 용어로서, 최근 새로운 생체인식 정보로 부각되고 있다[1]. 지문인식, 홍채인식 등의 기존의 생체인식 기술은 정확한 정보 추출을 위해서는 영상의 해상도가 매우 높아야하는 단점을 가지고 있다. 그러나 게이트 인식의 경우에는 영상의 해상도가 높지 않은 조건에서도 인식에 필요한 정보를 추출할 수 있는 장점이 있다.

게이트가 사람을 분별할 수 있음은 이미 증명되었으나, 비전 기술을 이용한 영상 기반 게이트 인식의 성능은 그리 높지 못하다. 그 이유는 실제 물리적인 게이트 정보와 영상을 기반으로 추출한 게이트 정보 사이에 많은 차이를 보이고 있기 때문이며, 아직까지 3차원 정보와 투영된 영상을 기반으로 추정된 정보 사이의 차이를 최소화 할 수 있는 효과적인 알고리즘이 제시되지 못하고 있다.

비전 기술을 이용한 게이트 인식에서 해결해야 될 가장 큰 문제는 영상을 기반으로 추출한 게이트 정보가 카메라의 시점에 종속적으로 변한다는 점이다. 카메라의 시점에 따른 인식을 저하 문제에 대해서는 게이트를 다양한 시점에서부터 획득하고 이를 통하여 정확한 3차원 정보를 예측하는 것이 가장 좋은 해결 방안이라 할 수 있다. 실제 카네기멜론 대학의 게이트 데이터베이스(MoBo)는 시점에 변화를 준 6대의 카메라로 게이트를 획득한 것으로 3차원 게이트 인식에 활용될 수 있다[2]. 그러나 이러한 경우는 사람의 걸음걸이를 획득하기 위하여 동기가 맞춰져 있는 여러 대의 카메라를 사용하여야 하고 게이트의 3차원 정보 추출을 위해 부가적인 연산이 필요한 단점을 가지고 있으며, 실제 응용에는 적합하지 못하다.

단일 카메라로부터 획득한 게이트 영상 시퀀스로부터 시점 변화에 비종속적인 정보를 추출하기 위한 연구도 다양하게 진행되고 있다[3][4]. [3]의 경우에

는 단일 시점의 영상 시퀀스로부터 실제 보폭 (Stride length)과 걸음걸이 주기(Cadence) 및 키 (Height) 추정을 통해 카메라 시점에 무관한 게이트 인식을 시도하였다. 그러나 정확한 카메라 보정이 되어있어야 하며, 사람의 위치와 카메라간의 3차원 깊이(Depth) 정보 등을 정확히 알고 있어야 하는 단점이 있다. [4]의 경우에는 걷는 방향과 카메라가 이루는 각을 영상 시퀀스로부터 추정하여 시점을 보정한 영상을 구성하는 방법으로 문제 해결을 시도하였다. 그러나 이 경우에도 역시 카메라 파라미터를 정확히 알고 있어야 하기 때문에 카메라 파라미터가 변하는 경우에 적용할 수 없는 단점이 있다. 본 논문에서는 카메라 보정이 되지 않은 카메라로부터 획득한 영상 시퀀스내의 정보만을 이용하여 평면 호모 그래피(Homography)를 추정하여 투영된 게이트를 보정함으로써, 키 및 보폭 등과 관련된 게이트 특징 정보를 구하는데 발생하는 오류를 줄일 수 있는 방법을 제안한다.

2. 카메라 측면 시점 영상 재구성

게이트 인식에 필요한 정보를 얻기 위해서는 카메라의 방향과 수직인 방향으로 걸어가는 영상을 사용하는 것이 가장 좋다. 기존의 게이트 인식에 사용된 데이터들을 살펴보면 대부분이 이러한 측면 시점 (Canonical-view)의 영상들을 사용하고 있다[5][6]. 그러나 카메라의 시점이 변할 경우 추출된 정보는 측면 시점에서의 정보와 많은 차이를 보인다.

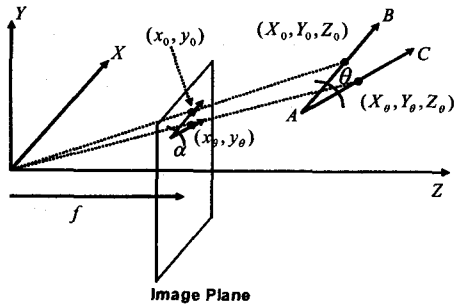


그림 1. 영상 기하학

그림 1은 3차원 좌표가 영상으로 투영되는 관계를 도식화한 그림으로 AB 방향이 카메라의 방향과 수직인 방향을 나타내고 이 방향으로 걸어갔을 때의 시점이 측면 시점이 된다. AC는 측면 시점과 특정 각도를 유지하여 걸어가는 방향을 나타낸 것으로 두 평면이 이루는 각 θ 를 정확히 알 경우 임의의 시점

영상을 측면 시점 영상으로 재구성할 수 있다.

2.1 Calibrated Case

사람이 직선으로 걷는다고 가정했을 때 실루엣이 이루는 평면과 카메라와 수직인 평면이 이루는 3차원 각 θ 를 알 수 없기 때문에 영상에 투영된 각 α 를 통하여 θ 를 추정하게 된다. 영상 평면상의 두 라인이 이루는 각 (α)과 3차원 상의 라인이 이루는 각 (θ)간의 관계는 식(1)과 같이 표현될 수 있다[4].

$$\tan(\theta) = \frac{1}{K} \tan(\alpha), \quad K = -\frac{k}{f} \quad (1)$$

여기서 k 는 영상에서의 y좌표 값을 사용할 수 있으며, f 는 카메라 초점거리를 나타낸다. 3차원 상의 두 점간의 관계는 식(2)와 같다. 이 관계로부터 측면 시점 영상을 재구성하면 측면 시점 영상에서의 좌표 (x_0, y_0) 는 식(3)과 같이 계산될 수 있다.

$$\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} = R(\theta) \begin{bmatrix} X_\theta \\ Y_\theta \\ Z_\theta \end{bmatrix}, \quad R(\theta) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \quad (2)$$

$$x_0 = f \frac{x_\theta \cos(\theta) - f \sin(\theta)}{f \cos(\theta) + x_\theta \sin(\theta)} \quad (3)$$

$$y_0 = f \frac{y_\theta}{-x_\theta \sin(\theta) + f \cos(\theta)}$$

2.2 Uncalibrated case

사람의 걸음걸이에 대한 2차원 투영 영상은 걸음걸이의 방향성을 나타내준다. 그림 2는 카메라 시점과 수직인 방향과 수직이 아닌 임의의 방향으로 걷는 경우에 대한 각각의 누적 실루엣 영상이다.

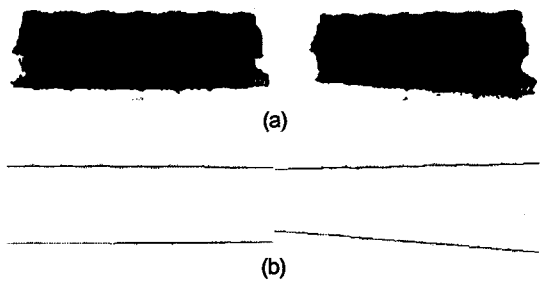


그림 2. 방향에 따른 누적 실루엣(a)과 직선 근사화(b) 카메라 시점과 수직방향(좌), 임의의 방향(우)

그림에서 보는 바와 같이 걸음걸이를 하나의 평면

으로 근사화 할 수 있으며, 본 논문에서는 이 평면을 게이트 평면이라 칭한다. 각 프레임에서의 머리와 발의 위치를 추적하고 추적된 각 점들을 그림 2의 (b)와 같이 직선으로 근사화 했을 때, 측면 영상은 두 직선이 평행에 가까운 특성을 나타냄을 알 수 있다. 이러한 특성으로부터 임의의 방향으로 걷는 게이트 영상에서의 두 직선을 평행하게 만드는 호모그래피는 곧 임의의 방향에 대한 게이트 평면을 측면 시점에서의 게이트 평면으로 변화시키는 평면 호모그래피라 할 수 있으며, 이를 이용하여 측면 시점 영상으로 재구성할 수 있다.

3. 제안하는 평면 호모그래피 추정

먼저 각 실루엣 영상에서 사람의 머리와 발의 위치를 추적하고 추적된 점들을 하나의 직선으로 근사화 한다. n개의 프레임에서 i번째 추적한 점의 좌표를 (x_i, y_i) 라 할 때 모든 점을 지나는 직선의 방정식에 대한 행렬식은 식(4)와 같이 표현된다.

$$\begin{bmatrix} x_1 & y_1 & 1 \\ \dots & \dots & \dots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \quad (4)$$

추적된 점을 근사화 하는 직선의 방정식은 식(5)와 같이 A행렬에 대한 SVD(Singular Value Decomposition)를 통해 구할 수 있다. 여기서 직선 파라미터 $[a, b, c]^T$ 는 식(6)과 같이 가장 작은 특이값을 가지는 V의 열벡터(Column Vector)가 된다.

$$A = \begin{bmatrix} x_1 & y_1 & 1 \\ \dots & \dots & \dots \\ x_n & y_n & 1 \end{bmatrix} = UDV^T \quad (5)$$

$$[a \ b \ c]^T = [v_{0s} \ v_{1s} \ v_{2s}]^T \quad (6)$$

본 논문에서는 에피폴라 라인을 수평 방향에 평행하게 만드는 방식[7]을 이용하여 게이트 평면 호모그래피를 추정하였다. 머리와 다리의 위치를 추적하여 근사화한 직선을 각각 l_1, l_2 라고 할 때 두 직선의 교차점 p_x 는 식(7)과 같이 구해할 수 있다. 이 때 p_x 를 무한대의 점, p_∞ 로 보내는 호모그래피 H는 식(8)과 같이 표현할 수 있으며, 여기서의 H가 곧 추정된 게이트 평면 호모그래피가 된다.

$$p_x = l_1 \times l_2 \quad (7)$$

$$p_\infty = Hp_x \quad (8)$$

$$H = GRT \quad (9)$$

식(9)에서 T는 영상의 중심 좌표를 원점으로 이동시키는 이동변환이고, R은 p_x 를 x축 상의 점 $[f, 0, 1]^T$ 로 회전시키는 회전변환이다. G는 점 $[f, 0, 1]^T$ 를 무한점 $[1, 0, 0]^T$ 로 사상하는 변환이다. 이렇게 추정된 평면 호모그래피에 의해 재구성된 측면 시점 영상의 누적 실루엣을 그림 3에 보였다.

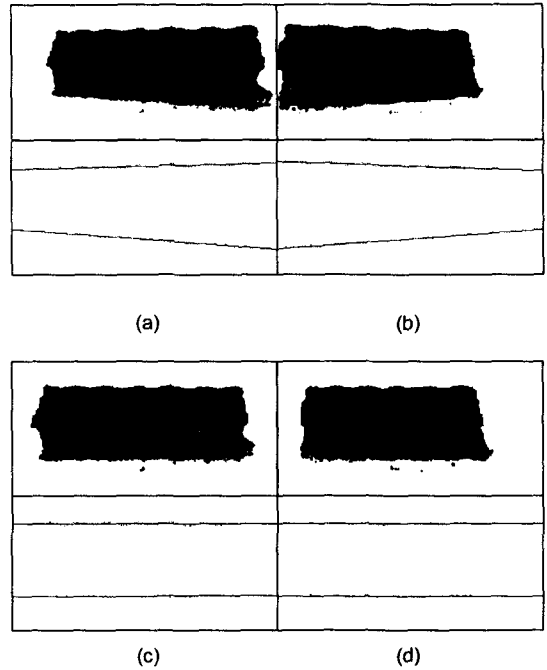


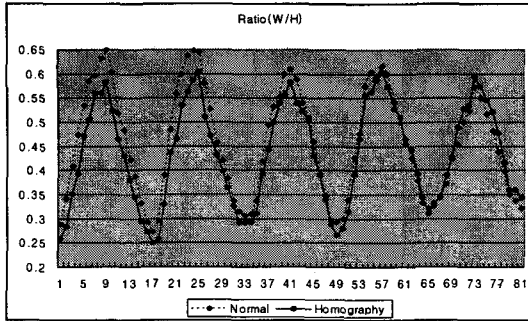
그림 3. 임의의 방향에 대한 누적 실루엣(a)(b)과 재구성된 측면 시점의 누적 실루엣(c)(d)

4. 실험 결과

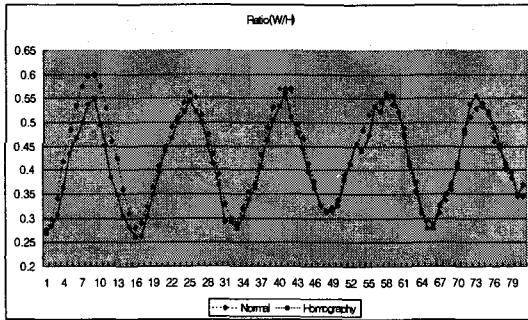
게이트 인식에서 키와 보폭은 개인을 구별할 수 있는 중요한 특징이 된다[3]. 이 정보는 카메라와 사람간의 거리에 따라 달라지기 때문에 정규화과정이 필요하다. 이는 실루엣 영역의 크기를 정규화 함으로써 문제를 해결할 수 있다. 그러나 그림 4의 (a)와 같이 측면 시점 영상이 아닌 경우에는 키와 보폭의 변화 비율이 게이트 평면의 각(θ)에 영향을 받는다. 이러한 경우 본 논문에서 제안한 측면 시점 영상 재구성 방법을 적용하여 문제를 해결할 수 있다. 그림 4의 (b)에서 보여 지듯이 제안한 시점 보정 방법은 실루엣의 두께가 왜곡되는 단점은 있으나, 키가 정규화 되고, 보폭도 거의 일정하게 유지된다. 특히 모든 상태에서의 키를 동일하게 만드는 일반적인 정규화와 달리 각 걸음의 상태에 따라 키가 변하는 특성을 그대로 유지하고 있다.



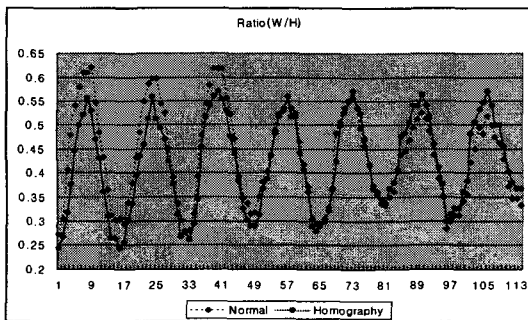
(a) 임의의 시점 영상 (b) 재구성된 측면 시점 영상
그림 4. 측면 시점 영상 재구성 결과



(a) 사람 1



(b) 사람 2



(c) 사람 3

그림 5. 사람에 따른 높이와 폭의 비율(W/H) 변화

그림 5는 측면 시점이 아닌 임의의 방향으로 걷는 영상에 대한 실루엣 영역의 변화를 보여주는 그래프로써, x축은 프레임 번호를, y축은 실루엣 영역의 폭과 높이의 비를 나타낸다. 이는 카메라로부터 멀

어지는 방향으로 걷는 경우로, 실선은 제안한 방법에 의해 보정된 결과이고, 점선은 보정되지 않은 결과이다. 그림에서 보듯이 시점을 보정해 주지 않으면, 카메라에서 멀어짐에 따라 폭이 높이보다 상대적으로 크게 감소하기 때문에 비율(W/H)이 점차 작아지는 문제가 있다. 이에 비해 제안한 시점 보정 방법을 적용했을 때에는 이러한 문제가 개선된다. 특히 최대값이 거의 일정하게 유지되고 있는데, 이 최대값은 정규화된 보폭을 의미하므로 일정한 값을 유지하는 것이 매우 중요하다.

5. 결론

본 논문에서는 간단한 연산을 통하여 게이트 평면에 대한 평면 호모그래피를 추정하고, 이 평면 호모그래피를 사용하여 측면 시점 영상으로 재구성하는 방법을 제안하였다. 제안한 방법은 게이트 인식의 시점 종속적인 단점을 보완할 수 있는 방법으로써, 특히 정규화된 보폭이 일정한 값을 유지함으로써, 시점에 무관한 보폭 정보를 얻을 수 있음을 확인하였다.

본 논문에서 제안한 방법은 실루엣의 두께가 왜곡되는 단점이 존재한다. 이를 개선하기 위한 추가적인 연구가 필요하다. 또한 보정된 정보를 이용한 인식에 관한 연구가 진행되어야 할 것이다.

참고문헌

- [1] D. Cunado, M. Nixon, J. Carter, "Gait Extraction and Description by Evidence Gathering.", Proceeding of 2nd Int. Conf. on Audio- and Video-Based Biometric Person Authentication, 1999.
- [2] R. Gross, J. Shi, "The CMU Motion of Body(MoBo) Database", Tech. report CMU-RI-TR-01-08, Robotics Institute, CMU, June 2001.
- [3] Chiraz BenAbdelkader, Ross Cutler, "View-invariant Estimation of Height and Stride for Gait Recognition", In Workshop on Biometric Authentication ECCV, 2002.
- [4] Amit K. Roy Chowdhury, Amit Kale, Rama Chellappa, "Video Synthesis of Arbitrary Views for Approximately Planar Scenes", ICASSP, 2003.
- [5] J. J. Little and J. E. Boyd, "Recognizing people by their gait: the shape of motion", Videre, 1(2), 1998, at <http://mitpress.mit.edu/e-journals/VIDE/001/v12.html>
- [6] L. Lee, W.E.L. Grimson, "Gait Analysis for recognition and classification", ICASSP, May 2002.
- [7] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000
- [8] D. A. Winter, *The Biomechanics and motor Control of Human Movement, 2nd Eds.*, John Wiley & Sons, 1990.