

능동 카메라에서 특징점을 이용한 객체 추적 알고리즘

김재호*, 김신형*, 황태현**, 주인학**, 장종환*

*배재대학교 정보통신공학과

**한국전자통신연구원 공간정보기술센터 실감 GIS 연구팀

e-mail : *{legnakcalb, zeros, jangjw}@mail.pcu.ac.kr, **{hth63339, ihjoo}@etri.re.kr

An Object Tracking Algorithm Using Feature Point of Active Camera

Jae-Ho Kim*, Shin-Hyung Kim*, Tae-Hyun Hwang**, In-Hak Joo**, Jong-Whan Jang*

*Dept. of Information and Communication Engineering, Pai-Chai University

**Tangible GIS Team, Spatial Information Technology Center, ETRI

요 약

본 논문에서는 비디오 지리정보시스템 구축을 위해 특징점을 이용한 객체 추적 알고리즘을 제안한다. 기존의 특징점을 이용한 객체 추적 알고리즘은 카메라의 움직임이 배제된, 즉 배경이 고정된 영상에서 객체의 특성을 이용한 방식으로, 운행중인 차량에 부착된 카메라에서 획득한 영상에서는 배경이 움직이므로 객체를 정확하게 추적하지 못하는 문제점이 있다. 본 논문에서는 이러한 문제점을 개선하기 위해 연속된 카메라 영상에서 모션벡터를 이용해 시간에 따른 객체의 변화율을 유도하고 유도된 변화율을 이용하여 객체 추적 알고리즘을 제안한다.

1. 서론

비디오 지리정보시스템(VideoGIS : Video Geographic Information System)이란 지리공간정보(Geographic Information Space)를 단순히 2D/3D 수치지도로써 보여주는 것이 아니라, 비디오 정보를 기반으로 하여 직접 사용자와 상호작용이 가능하고 공간 데이터를 분석하고 가공하는 시스템을 말한다.[1] 따라서 VideoGIS 를 통하여 사용자들은 기존의 GIS 로부터 얻을 수 있었던 정보뿐만 아니라 시설물이 어떠한 모습으로 존재하고 있는지에 대한 현실적이고 시각적인 정보를 얻을 수 있다.[2]

이러한 VideoGIS 를 구축하기 위해서는 카메라가 장착된 차량을 통해 도로를 주행하며 도로시설물 및 건물 등의 비디오를 획득하고, 이렇게 얻어진 비디오 데이터에서 시설물의 지리 위치정보를 추출하고, 추출된 데이터를 관리하는 데이터베이스 시스템을 구축해야 한다.[3] 그러기 위해 비디오 데이터의 취득 위치를 측정하고, 비디오 데이터에서의 시설물을 추출하며, 추출된 시설물의 지리 위치정보를 계산하는 작업들이 필요하다. 이 중 비디오 데이터에서의 시설물을 자동으로 추출하기 위해서는 비디오내의 객체 추출 및 추적에 관한 연구가 필수적이다.

본 논문에서는 동영상에서 모션 벡터 정보를 이용하여 카메라의 움직임에 따른 객체의 변화율을 유도

하고, 객체의 변화율과 특징 점을 이용하여 배경에 영향을 받지 않는 개선된 객체 추적 알고리즘을 제안한다.

논문의 구성은 2 장에서 기존의 객체 추적 알고리즘에 대해서 설명하고, 3 장에서는 카메라의 움직임에 따른 영상 내 객체의 변화율을 유도한다. 4 장에서는 변화하는 배경영역에 영향을 받지 않고 객체를 추적할 수 있는 개선된 객체 추적 알고리즘을 제안한다. 5 장은 제안한 알고리즘을 실험을 통해 성능을 분석하고, 6 장에서 결론을 기술한다.

2. 기존의 객체추적 알고리즘

객체 추적 방법[4]에는 크게 두 가지 기법으로 분류될 수 있다. 그 중 하나는 프레임 내 픽셀들의 밝기 변화나 색깔의 특징을 이용한 공간적 객체 추적 방법 [5]인데 이 방법에는 객체내의 픽셀 값의 특성을 분석하여 추적하는 방법과 픽셀의 색차 특성을 분석하여 추적하는 방법이 있다. 픽셀 값의 특성을 이용한 방법에는 블록 매칭 알고리즘(BMAs : Block Matching Algorithms)[6]으로 서로 매칭되는 각 블록간의 픽셀 값의 상관도가 최대가 되는 블록을 찾는 기법이다. 이 기법은 배경이 움직이는 영상에 적용할 수 있지만, 객체의 모양이 크게 변하는 경우에는 검색영역 내에 객체를 포함하고 있지 않아 매칭되지 않는 문제가 발생한다.

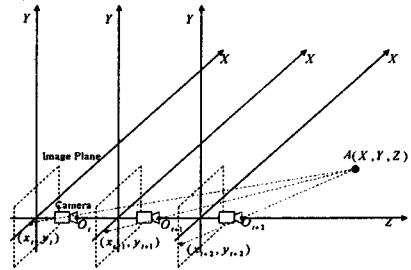
픽셀의 색차 정보를 분석하여 추적하는 방법은 컬러 모델(Color Model)을 이용하는데 HSI(Hue, Saturation and Intensity) 컬러공간(Color Space)을 이용한 기법[7][8]이 있다. 이는 특정 범위에 존재하는 색상의 값을 검색하기 용이하다는 점을 이용하여 도로에 존재하는 특정 표지판이나 자동차의 번호판 추적에 적용할 수 있지만 배경 영역과 색상이 유사한 가로등이나 건물 등을 추적하기에는 적용하기 어려운 문제가 있다.

다음 기법으로는 연속된 영상에서 프레임간의 특징, 즉 시간적 변화에 따른 영상의 특징을 분석하여 객체를 추적하는 시간적 객체 추적 방법[9]이다. 이 방법에는 연속된 영상 내에서 객체가 이동하면, 객체를 구성하는 바탕(Texture)에 의해 움직이기 전과 움직인 후의 픽셀의 밝기(Intensity)가 변화하는 특징을 이용하여 밝기 변화가 없는 영역을 비 관심영역(배경 영역), 밝기 변화가 있는 영역을 관심영역(객체)으로 세분화하여 특정 객체의 움직임을 추정하는 방법과 이전 영상에서 객체의 움직임 특징을 분석하여 현재 영상에서 가장 유사한 움직임 특징을 찾아내 객체의 움직임을 추정하는 방법이 있다. 이러한 방법들 또한 운행중인 차량에 부착된 카메라에서 획득한 영상에서는 객체만이 아니라 배경 영역에 변화를 수반하기 때문에 관심 객체를 정확하게 추적하기 어렵다.

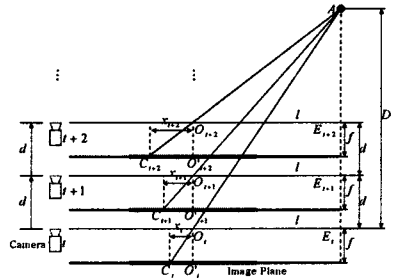
3. 능동적 카메라에 따른 영상 정보의 특징

차량주행을 통해서 획득한 영상정보는 카메라가 정적이 아닌 동적인 움직임을 지니고 있기 때문에 VideoGIS 를 구축하기 위한 관심영역(공간객체정보 : 건물, 시설물 등)만이 아니라 비 관심영역(사람, 차량, 도로 등)도 일정 비율로 변화한다. 따라서 능동적 카메라에 따른 영상 정보의 특징 분석이 필수적이다.

[그림 1]은 직진하는 차량에 부착된 카메라의 움직임에 의해 객체가 영상 평면(Image Plane)에 맺히는 상의 기하학적 구조를 나타내고 있다. X, Y, Z 는 실세계 좌표계(World Coordinates)에서 3 차원 좌표축을 나타내고, $A(X, Y, Z)$ 는 객체의 3 차원 좌표를 나타낸다. 여기서 차량주행에 따른 카메라의 이동은 직선구간에 대해서만 고려하였다. 직선구간에서 카메라의 움직임에 의해 좌표 (X, Y, Z) 에 위치한 객체 A 가 초점을 투영하여 맺힌 상은 픽셀 좌표 (x_t, y_t) 에서 (x_{t+1}, y_{t+1}) 로, 다시 (x_{t+1}, y_{t+1}) 에서 (x_{t+2}, y_{t+2}) 로 이동한다. 즉, 직진구간에서의 픽셀들의 움직임을 [그림 1]의 구조로 나타낼 수 있다. [그림 2]는 [그림 1]의 3 차원 구조를 픽셀 좌표계(Pixel Coordinates)인 y 축에서 바라본 기하학적 구조를 나타내는 것이다. t 는 현재 프레임의 시간을 나타내며, O 는 카메라의 초점(Focus)을, O' 는 영상평면의 중심, f 는 초점거리(Focal Length), d 는 카메라의 이동거리, A 는 관심객체, C 는 객체 A 가 카메라의 초점을 투영해 생긴 객체의 상, D 는 객체 A 와 카메라와의 전체거리의 미한다.



[그림 1] 직진하는 차량에 부착된 카메라의 3 차원적인 기하학적 구조



[그림 2] y 축에서 바라본 카메라의 기하학적 구조

카메라가 일정한 거리 d 만큼 이동한다고 가정한다면, 카메라와 객체와의 전체 거리 D 는 d 의 α 배로 표현할 수 있다. 여기서 객체 A 와 t 시점에서 카메라의 초점을 투영해 생긴 상 C_t 에 의해 생긴 선분들은 삼각형 $\Delta AO_t E_t$ 와 $\Delta O_t C_t O'$ 을 이루며, 이 두 삼각형은 서로 닮은꼴 삼각형이다. 따라서 닮은꼴 삼각형의 성질에 의해 대응하는 변의 길이의 비는 일정하므로, 두 닮은 삼각형의 대응하는 두 변의 길이의 비는 식 (1)과 같다.

$$\overline{AE_t} : \overline{O_t O'} = \overline{O_t E_t} : \overline{C_t O'} \quad (1)$$

$$\alpha d : f = l : x_t$$

또한 $t+1$ 시점에 위치한 카메라의 초점을 투영해 생긴 상 C_{t+1} 와 $t+2$ 시점에서의 C_{t+2} 도 닮은꼴 삼각형의 성질에 의해 식 (2)와 (3)과 같이 얻을 수 있다.

$$(\alpha - 1)d : f = l : x_{t+1} \quad (2)$$

$$(\alpha - 2)d : f = l : x_{t+2} \quad (3)$$

식 (1)과 (2), (3)을 통해 카메라가 $t+n$ 시점에 위치해 있을 때 생기는 대응 비는 식 (4)와 같이 정의할 수 있다.

$$(\alpha - n)d : f = l : x_{t+n} \quad (4)$$

다음 식 (1)과 식 (2)을 x_t, x_{t+1} 에 관해 풀면 식 (5)와 같이 되며, 식 (2)와 식 (3)에 대해서도 같은 방법으로 풀면 식 (6)과 같이 풀이된다.

$$x_{t+1} = \frac{\alpha}{\alpha - 1} x_t \quad (5)$$

$$x_{t+2} = \frac{\alpha - 1}{\alpha - 2} x_{t+1} \quad (6)$$

따라서 식 (5)와 (6)을 통해 $t+n-1$ 시점에 맺힌 상의 x 좌표가 $t+n$ 시점에서의 위치는 식 (7)을 이용해 일반화시킬 수 있다.

$$x_{t+n} = \frac{\alpha - n + 1}{\alpha - n} x_{t+n-1} \quad (7)$$

또한 일반화된 식 (7)을 이용하여 x 의 변화량을 구하면 식 (8)과 같이 풀이되며, y 의 변화량도 위와 같은 과정을 통해 식 (9)와 같이 일반화된 식으로 유도할 수 있다.

$$\Delta x_{t+n} = x_{t+n} - x_{t+n-1} = \frac{1}{\alpha - n} x_{t+n-1} \quad (8)$$

$$\Delta y_{t+n} = y_{t+n} - y_{t+n-1} = \frac{1}{\alpha - n} y_{t+n-1} \quad (9)$$

식 (8)과 식 (9)를 보면 알 수 있듯이 차량이 일정한 거리 d 만큼 이동한다면, 현재 프레임에 존재하는 객체의 위치 정보로 다음 프레임내의 이동한 객체의 위치 정보를 예측해 낼 수 있다. 즉, 능동적 카메라에서 획득한 연속적 영상내의 객체의 위치는 α 의 값으로 다음 프레임에서의 객체의 위치를 예측할 수 있으며, 이는 일정한 속도로 운행하는 차량에 부착된 능동적 카메라에 의한 영상내의 객체의 특징으로 정의할 수 있다.

4. 제안하는 새로운 객체 추적 알고리즘

본 논문에서는 능동적 카메라에 의한 영상내의 객체의 특징을 분석하여, 관심객체를 추적한다. 즉, 일정한 거리로 직진을 차량에 부착된 카메라로 획득한 연속된 영상에서 관심객체의 변화량은 현재 관심객체의 위치에 $\frac{1}{\alpha - n}$ 배 한다는 특징을 이용하여 다음 프레임내의 객체의 위치를 예측한다. 또한 이를 적용하기 위해 우리는 연속된 영상에서 일정 구간 동안에는 차량의 속도는 일정하다고 가정하였다.

여기서 본 논문에서 객체의 움직임 비율로 유도한 α 을 구하기 위해서는 객체와 카메라 사이의 전체거리 D 와 차량이 이동하는 거리 d 의 값을 알아야 한다. 그러나 영상만 주어진 상태에서 이러한 데이터를 수집하기에는 어려운 일이다. 따라서 α 을 구하는 방

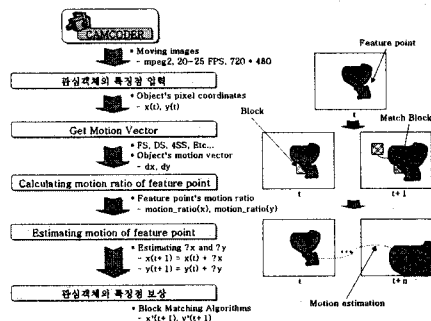
법은 이전 프레임에서의 객체 위치 x_{t+n-1} 와 현재 프레임에서의 객체위치 x_{t+n} 와 차이를 이용한다. 초기 특징점의 위치를 x_{t+n-1} 로 하고, 다음 프레임에서 이동한 x_{t+n-1} 의 모션벡터를 구해 식 (9)로 x_{t+n} 을 구한 다음 식 (10)을 이용하여 α 를 얻을 수 있다.

$$x_{t+n} = x_{t+n-1} + dx \quad (9)$$

$$\frac{\alpha - (n-1)}{\alpha - n} = \frac{x_{t+n}}{x_{t+n-1}} \quad (10)$$

α 는 연속된 영상에서 맨 처음 프레임과 두 번째 프레임을 이용해 한번만 구하고 다음 프레임에서 이동한 객체의 위치를 자동으로 예측한다. 또한 y 축에 대한 특성도 x 축에 대한 특성과 같으므로, 위에서 수행한 방식과 동일한 방식으로 y 축에 대한 α 를 유도한 후 y 에 대한 객체의 이동 위치를 예측한다. 제안된 객체 추적 알고리즘의 전체 구성도는 [그림 3]과 같다.

[그림 3]에서 보듯이 제안된 객체 추적 알고리즘은 5 단계의 구성으로 이루어져 있다. 첫 번째 과정은 사용자가 임의로 관심객체의 모서리를 특징점으로 설정한다. 두 번째 과정에서는 첫 번째 과정에서 설정한 특징점을 (x_t, y_t) 로 정하고, 다음 프레임에서 이동한 초기 특징점의 이동 위치를 구해 (x_{t+1}, y_{t+1}) 로 정한다. 다음 식 (9)를 이용하여 α 를 구한다. 세 번째로 두 번째 과정에서 유도한 α 를 가지고 식 (8)과 식 (9)를 이용하여 다음 프레임에서 이동한 (x_{t+n}, y_{t+n}) 의 위치를 예측한다. 네 번째 과정에서는 카메라의 흔들림에 의해 생기는 객체의 예외적인 움직임에 의한 문제를 해결하기 위해 예측한 (x_{t+n}, y_{t+n}) 을 중심으로 일정 영역을 블록 매칭 알고리즘을 이용하여 보정한다. 다음 일정한 구간마다 새로운 α 를 갱신하기 위해 반복횟수 n 이 최대 반복횟수 N 에 도달했는지 검사 후 도달하지 않았다면, 계속 이전 α 로 객체의 움직임을 예측하고, 도달하였다면 위의 두 번째 과정부터 다시 반복하여 객체를 추적한다.



[그림 3] 제안된 객체 추적 알고리즘 전체 구성도

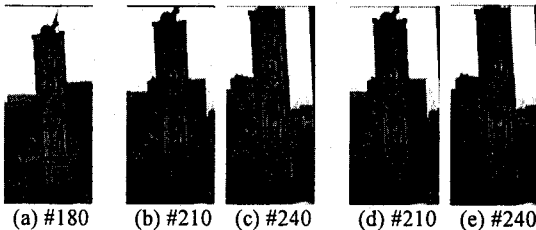
5. 실험 결과

본 연구에서 제안된 객체 추적 알고리즘을 이용하여 실험한 결과를 제시하고, 그 결과를 분석하였다. 실험에 사용된 테스트 영상은 차량에 부착된 카메라를 통해 획득하였으며, 차량의 속도는 시속 30~40km로 주행하였다. 영상의 크기는 640*480 이며, 실제 도로 주변의 시설물을 촬영한 영상이다. 또한 제안된 알고리즘과 기존 알고리즘의 성능비교를 위해 블록 매칭 기법만을 사용한 객체 추적 알고리즘과 비교하였으며, 실험 환경은 펜티엄 4 프로세서 2.2GHz 의 CPU 와 Windows 환경에서 MFC 를 사용하였다.

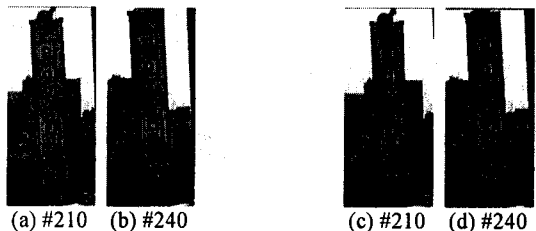
[그림 4]와 [그림 5]의 실험에 사용된 파라미터는 [표 1]과 같이 사용 되었으며, 각각 30 프레임씩 수행하였다.

[표 1] 실험에 사용된 파라미터

알고리즘	블록크기	검색범위		
BMA	9*9	실험 1	32*32	
		실험 2	64*64	
제안한 알고리즘	9*9	모션벡터	보정	
		실험 1	32*32 16*16	
		실험 2	64*64	16*16



[그림 4] 객체추적 실험 결과
(a) 초기 특징점 설정, (b)~(c) BMA 기법(검색범위 32*32), (d)~(e) 제안한 기법(검색범위 16*16)



[그림 5] 객체추적 실험 결과
(a)~(b) BMA 기법(검색영역 64*64), (c)~(d) 제안한 기법 (검색범위 16*16)

[표 2] 실험 수행시간

알고리즘	수행시간	
BMA	실험 1	104 초
	실험 2	439 초
제안한 알고리즘	실험 1	45 초
	실험 2	198 초

위의 그림에서 기존 알고리즘 보다 제안한 알고리즘이 더 정확하게 객체를 추적하는 것을 알 수 있는데, 이는 변화범위가 점점 커지기 때문에 검색범위를 벗어나 올바른 특징점을 찾을 수 없기 때문이다.

또한 [표 2]를 보면 알 수 있듯이 유사한 추적결과를 얻으면서, 기존 알고리즘보다 속도가 향상된 것을 알 수 있었다.

6. 결론

본 논문에서는 동영상에서 모션 벡터 정보를 이용하여 카메라의 움직임에 따른 객체의 변화율을 유도하고, 객체의 변화율과 특징 점을 이용하여 배경에 영향을 받지 않는 개선된 객체 추적 알고리즘을 제안하였다. 또한 유도한 객체의 변화율을 이용하여 객체의 이동경로를 예측하여 빠르고 정확하게 객체를 추적하였다. 실험에서 제안한 알고리즘은 객체의 이동경로를 예측하여 가장 유사한 특징점을 찾기 때문에, 검색범위가 넓지 않더라도 기존 알고리즘보다 성능이 우수함을 입증하였다.

향후 연구 과제로는 본 논문에서 유도한 객체의 변화율은 직선적인 차량의 움직임만 고려하였기 때문에, 차량이 우회전이나 좌회전 시 객체를 올바르게 찾지 못하는 단점이 있으므로 차량의 코너링 경우에 객체의 변화율에 대한 연구가 계속되어야 한다고 사료된다.

참고문헌

- [1] 유재준, 주인학, 남광우, 이종훈, "The Design and Implementation of A Video Geographic Information System," 한국정보과학회 2002 추계학술대회, 1 권, pp274-276, Oct. 2002.
- [2] Navarrete, T., "VideoGIS:Combining Video and Geographical Information," Research Report, Pompeu Fabra Univ., Dept. of Computer Science and Communication. 2001.
- [3] Manandhar, D., Shibasaki, R., "VehicleBorne Laser Mapping System for 3-D Urban GIS Databases," CUPUM 2001, July. 2001.
- [4] F.Marques and C. Molina, "Object tracking for content-based functionalities," in SPIE Vis. Commun. Image Processing, vol 3024, San Jose, CA, Feb. 1997, pp. 190-199.
- [5] L. Vincent and P. Soille, "Watersheds in digital spaces : an efficient algorithm based immersion simulations," IEEE Trans. on PAMI, Vol.13, No.6, pp.583-598, June 1991.
- [6] T.Koga, K.Ilimuma, A.Hirano, Y.Iijima, and T.Ishiguro, "Motion-compensated interframe coding for video conferencing," in Proc. NTC 81, New Orleans, LA, pp. C9.6.1-C9.6.5, Nov./Dec. 1981.
- [7] Tom Williams, "HIS conversion brings true color image processing to life," Computer Design, Jan. 1988, pp.28-29
- [8] F.Perez and C. Koch, "Towards Color Image Segmentation in Analog VLSI: Algorithms and Hardware," Int. J. Computer Vision, Vol.12, No.1, pp.17-42, 1994.
- [9] Luis Torres and Murat Kunt, Video Coding, pp.79-124, Academic Pub., 1996.