

Mellin Transform 에 기반한 파노라믹 영상 생성을 위한 이미지 접합

이지현*, 송복득, 윤태수, 양황규
*동서대학교 인터넷 공학부
e-mail : winter42@dreamwiz.com

Image Registration for Panoramic Image Using Mellin Transform

Ji-Hyun Lee*, Bok-Deuk Song, Tae-Soo Yun, Hwang-Kyu Yang
*Division of Internet Engineering, Dong-Seo University

요 약

본 논문은 카메라를 수평이동시켜 얻어진 영상을 Mellin Transform 을 이용하여 파노라마 이미지를 구성하는 방법을 제안한다. 일반적으로 카메라로부터 얻어진 동영상은 각 프레임 사이에 많은 시간적, 공간적 정보가 중복되어 있다. 이러한 중복 정보는 모자이크 방법 중의 하나인 파노라믹 기법을 사용하여 줄일 수 있다. 지금까지의 제안된 이미지의 중복성을 찾는 알고리즘들은 노이즈에 지나치게 민감하거나 중복성을 계산하는 데 있어 시간이 많이 소모되는 단점이 있었다. 따라서 본 논문에서는 Mellin Transform 을 사용하여 노이즈에 덜 민감하고 빠른 시간에 이미지의 중복 정보를 찾아내는 방법을 제안한다. Mellin Transform 을 이용하여 생성된 파노라마 영상은 이미지의 이동정보를 쉽게 구할 수 있어 동영상에서 움직이는 물체를 추적하거나 추정이 가능하여 물체 tracking 영역에 응용될 수 있을 것으로 보이며 또한 앞으로 현실감 있는 가상현실의 배경 생성에 도움이 될 것으로 생각된다.

1. 서 론

파노라마 이미지는 처음에 특별한 용도의 값비싼 하드웨어로 만들어 내는 전통적인 가상현실을 대신하여 제안되었다. 파노라마 이미지란 자신을 둘러싸고 있는 주위의 원통형 이미지로 카메라로부터 얻은 여러장의 영상을 하나의 넓은 영상으로 정합하는 것이다[1]. 파노라마 영상은 특별한 하드웨어가 필요 없이 실세계 환경으로 부터 표현되므로 기존의 3 차원 그래픽 기반으로 만들어진 가상현실 보다 실시간 표현이 가능하며 좀더 현실감 있게 표현할 수 있는 장점이 있다.

카메라로부터 얻은 영상 또는 동영상은 시간적인 정보를 표현할 수 있고 연속적으로 변화하는 정보를 표현할 수 있다는 장점이 있으나 이러한 점은 각 프레임간에 중복정보를 포함하고 있어 전체적으로는 정보의 양이 증가하는 단점이 있다. 따라서 파노라마 이미지는 각 프레임간의 중복 정보를 찾아서 하나의 큰 이미지로 표현하여 정보의 양을 줄일 수 있

다. 이러한 영상간의 중복 정보를 찾기 위한 연구들이 많이 제안되었다. 관련 연구로는 이미지 매칭 기술, moment invariants 와 푸리에 변환 등이 있다.

이 중 각 프레임간의 중복정보를 찾는 데 이용되는 이미지 매칭은 장면에서의 변화, 물체 행동의 추정, 목표의 위치, 물체를 식별하거나 이미지의 다른 형태로부터 정보를 구하는데 사용된다. 매칭 알고리즘은 노이즈가 많은 상황이나 같은 물체가 다른 강도, 다른 방향, 크기일 때 또는 물체가 한 이미지에서 부분적으로 불완전할 때 정교해야 한다[2].

이미지 매칭은 상호상관함수(cross-correlation) 기술을 기초로 한다. 그러나 자연 이미지에 대한 자기상관함수(auto-correlation)는 어느 정도 먼 거리에서 값을 가지므로 두 이미지 사이에 상호상관함수(cross-correlation)의 최대값의 차이가 커서 노이즈의 위치를 찾기가 어렵다[4]. 지금까지 상호상관함수(cross-correlation)에 관련된 여러 연구들이 제안되었으나 이러한 기술들은 두 이미지의 이동 정보에 대한 철저한 계산이 요구된다. 그러한 철저한

계산은 두 이미지간에 이동뿐 아니라 회전각, 크기가 포함될 때 시간이 많이 소모된다.

이미지 중복 정보를 찾기위한 또다른 기술은 moment invariants 에 기반한다[5]. Moment 는 노이즈에 매우 민감하기 때문에 대부분의 알고리즘은 이진화 이미지나 이진화된 그레이 이미지에 적용되는 한계점이 있다.

푸리에 변환의 위상을 이용한 방법은[7] 많은 연구에서 이미지 프로세싱 시스템에 이미지 매칭에 대한 위상 정보를 사용하기 위해 이루어졌다. 이 기술들의 장점은 높은 식별력, 효율적인 계산능력 그리고 노이즈에 대해 강하다는 것이다. 그러나 이 기술은 회전과 크기 변화가 작고 무시할 수 있을 때만 효과적이라는 문제점이 있다[8].

이러한 문제점을 해결하기 위해서 본 논문에서는 이미지의 중복성에 대한 정보를 효과적으로 찾아서 파노라마 영상을 생성하기 위해서 Mellin Transform 을 이용하고자 한다. 이에 대한 전체적인 흐름도는 그림 1 과 같다. 본 연구에서는 이미지의 수평이동에 중점을 두며 회전된 이미지에 대한 점은 제외되었다.

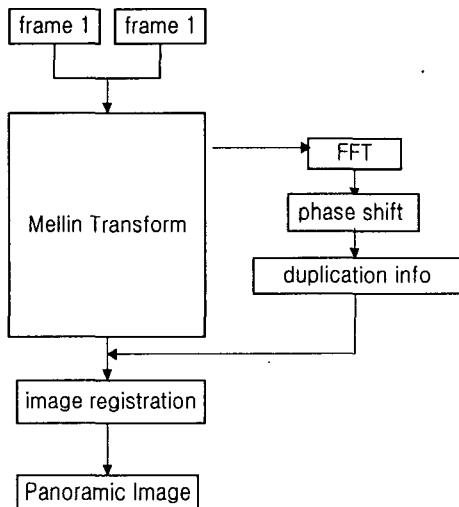


그림 1. 전체 구성도

2. 멜린 변환(Mellin Transform)

Mellin Transform 은 Fourier Transform 의 속성을 이용하여 주파수 공간에서 이미지의 이동성, 회전, 축소 및 확대, 위상 상관관계등의 정보를 추출하는 변환이다. 따라서 Mellin transform 은 위상 관계와 Fourier 속성에 기초를 둔다[3].

연속된 한 쌍의 이미지 $\{I_1, I_2\}$ 사이에서는 phase correlation 과 관계된 이동 모션인 (x_0, y_0) 를 알 수 있다. 즉, 두 영상 $I_1(x, y)$ 와 $I_2(x, y)$ 는 (x_0, y_0) 의 차이를 가지므로 이 정보를 멜린 변환을 이용하여 이미지의 이동정보를 추출할 수 있다.

$$I_2(x, y) = I_1(x - x_0, y - y_0) \quad (1)$$

각 이미지 I_1, I_2 에 Fourier transform 을 적용하여 두 영상의 phase shift 정보를 찾을 수 있다.

$$F_2(\omega_x, \omega_y) = e^{-j2\pi(\omega_x x_0 + \omega_y y_0)} F_1(\omega_x, \omega_y) \quad (2)$$

따라서, 두 영상의 phase shift 인 $e^{-j2\pi(\omega_x x_0 + \omega_y y_0)}$ 는 다음의 값을 가진다.

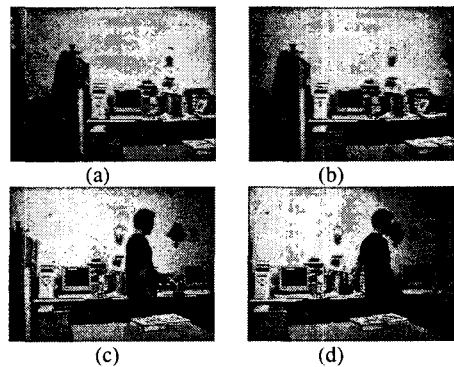
$$e^{-j2\pi(\omega_x x_0 + \omega_y y_0)} = \frac{F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)}{|F_1(\omega_x, \omega_y) F_2(\omega_x, \omega_y)|} \quad (3)$$

여기서 F_2^* 는 F_2 의 complex conjugate 이다.

이 phase shift 를 다시 역 푸리에 변환을 적용하면 원본 이미지에서 이동된 움직임의 양만큼 정확하게 delta function 으로 나타난다. 이것이 두 영상 사이의 평행이동 성분이다. 그러나 만약 실제로 이동이 많지 않다면 delta function 의 강도가 약할 수도 있다.

그림 2 는 이미지 접합시 나타나는 이미지의 수평이동에 대한 정보를 나타낸다. (b)의 이미지는 (a)의 이미지 위치에서 수평방향으로 이동한 이미지이다. 그에 반해 (c)의 경우는 카메라가 수평으로 이동하는 것과 동시에 (a)에는 없던 물체가 나타났을 경우이며 (d)는 (c)의 물체가 카메라 이동과 같은 방향으로 이동한 이미지이다.

(e)의 그래프는 (a)와 (b)의 이미지를 접합했을 경우의 delta function 이 나타나는 그래프이다. 이미지가 이동한 위치에서 delta function 이 나타난다. (f)의 그래프는 (a)와 (c)를 접합했을 경우이다. (c)는 (b)에 비해 수평이동의 양이 크다. 그 차이만큼 delta function 이 나타나는 위치가 달라지며 움직이는 물체가 나타남에 따라 그래프의 변화량이 많이 나타난다. (g)는 (a)와 (c)의 이미지를 접합한 후에 (d) 이미지를 다시 접합한 경우이다. 이 그래프에서는 (f)와 같은 모양의 그래프가 보인다. 이것은 두 장의 이미지를 접합한 후 접합된 이미지에 또 다른 이미지를 접합하더라도 처음 접합하는 이미지의 delta function 정보가 남아 있다는 것을 알 수 있다.



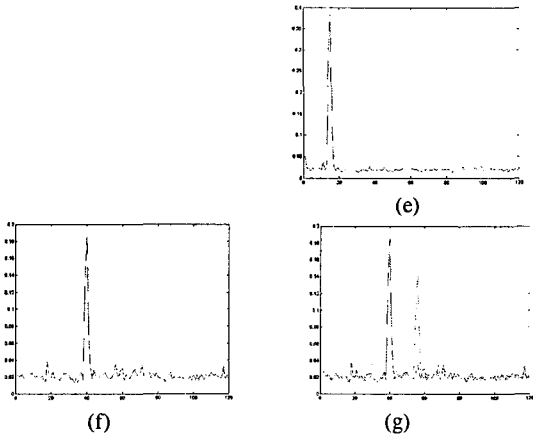


그림 2. Mellin Transform 실행후 나타나는 평행이동 성분

3. 이미지 접합

이미지를 접합하기 위해서는 접합을 위한 이미지의 중심축을 찾아야 한다. 이때 전체 이미지를 검색하여 중심축을 찾을 경우는 시간이 많이 걸리기 때문에 실험 시 나타난 delta function 을 이용하여 중심축을 찾기위한 범위를 축소하였다. 중심축은 delta function 이 나타난 범위를 중심으로 하여 두 영상을 비교하여 가장 작은 오차를 가지는 지점을 수식 (4)를 이용하여 찾아낸 후 이미지를 접합하였다.

$$\min E_{mse} = \min \left(\frac{1}{M} \sum [I_2(x_i, y_j) - I_1(x_i, y_j)]^2 \right),$$

$$\alpha \leq i \leq \beta, \quad 1 \leq j \leq M \quad (4)$$

여기서 i 는 delta function 이 나타나는 범위이며 M 은 이미지의 높이이다.

그림 3 은 이미지 접합 과정을 나타내고 있다.

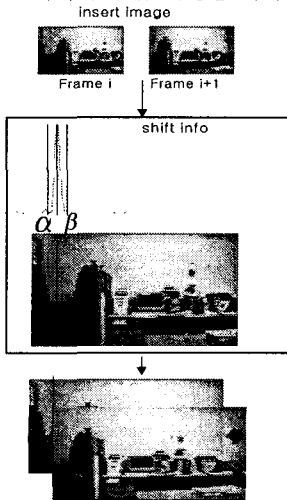


그림 3. 이미지 접합 과정

4. 실험 결과

본 논문에서 제안한 방법을 검증하기 위해서 연구실의 모습을 기본으로 하여 파노라마 영상을 제작하였다. 실험영상은 X 축 평행이동성분이 있으며 영상의 회전은 존재하지 않는다. 총 프레임 수는 20 프레임이며 카메라를 수평이동 시켜 이미지를 입력받았다. 그림 3 은 실험을 위해 제작한 카메라 시스템의 프로토타입이다.

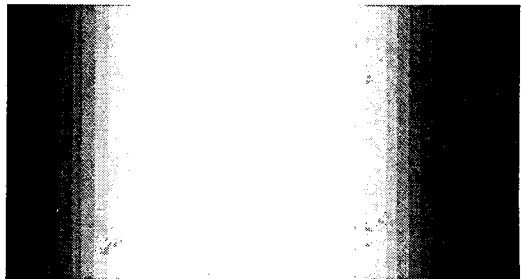


그림 3. 카메라 프로토타입

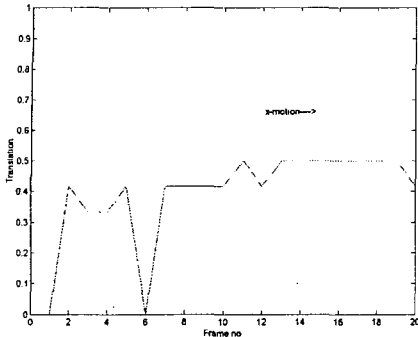
그림 4 에서 (a)는 실험실 내 실제 영상으로 생성된 파노라마 영상이며 (b)는 파노라마 생성시 각 프레임 간의 중복되는 정도를 나타낸다. 가장 많이 중복되는 부분은 명암도가 밝게 나타난다. (c)는 (a)의 파노라마 생성시 평행이동 정보만을 따로 추출한 결과이다.



(a)



(b)



(c)

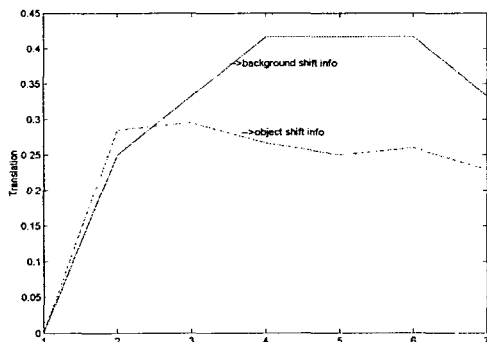
그림 4. 파노라마 영상 생성 결과

Mellin Transform 을 이용하여 파노라마 영상을 만들 때 배경의 이동정보 뿐 아니라 이미지 안의 움직이는 물체의 이동정보에 관한 정보도 알 수 있었다.

그림 5 는 카메라 이동과 동시에 이동하는 물체가 있을 때의 파노라마 영상 생성 시에 나타나는 영상이다. (a)는 파노라마 영상의 생성 결과이다. 물체가 이동한 흔적이 나타나고 있다. (b) 그래프는 카메라의 수평이동 정보와 움직이는 물체의 수평이동 정보를 나타낸다. 이 결과를 이용하면 움직이는 물체를 추출하여 좀더 정확한 파노라마 영상을 만들 수 있다.



(a)



(b)

그림 5. 이동하는 물체가 있을 경우 생성 결과

5. 결론

본 논문에서는 파노라마 영상 생성시에 이미지의 중복성을 좀 더 쉽게 찾고자 Mellin Transform 을 사용하였다. Mellin Transform 은 여러 이미지를 접합하는 과정에서 빠른 시간에 이미지가 중복되는 부분은 정확하게 찾아내었고 또한 이미지와 물체의 수평이동 정보를 쉽게 찾을 수 있었다. 그러나 Mellin Transform 을 이용하여 접합할 때 처음 이미지와 나중 이미지는 1/3 이상의 중복성을 가져야만 이미지가 접합이 되었다. 이는 카메라가 갑자기 이동하여 프레임간의 수평이동량이 커지게 되면 이미지가 접합되지 못하는 단점이 있다. 또한 수평이동 외에 이미지가 회전되었을 때 이미지의 중복성을 찾아서 접합하는 문제가 남아 있으며 카메라에 의한 노이즈를 줄이는 방법이 필요하다. 따라서 향후 이러한 문제점들을 보완하는 연구가 필요하며 Mellin Transform 을 이용하는 과정에서 나타나는 물체의 이동정보를 물체의 추적, 추정 등에 관한 연구에 이용하고자 한다.

참고문헌

- [1] Shmuel Peleg, "Omnistereo: Panoramic Stereo Imaging", *IEEE Trans. Pattern Analysis*, vol. 23, no. 3, pp.279-290, 2001
- [2] Qin-sheng Chen, Michel Deffrise, F. Deconinck "Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition", *IEEE Trans. Pattern Analysis*, vol. 16, no. 12, pp.1156-1168, Dec. 1994
- [3] J.Davis "Mosaics of Scenes with Moving Objects", *Computer Vision and Pattern Recognition Proceedings*, pp.354-360, 1998
- [4] W. K. Pratt, *Digital Image Processing*. New York: John Wiley & Sons, 1978, pp. 526-566
- [5] M.O.Freeman and B.E.A.Salch, "Moment invariants in the space and frequency domains", *J. Opt. Soc. Am. A.*, vol. 5, no. 7, pp. 1073-1084, 1988
- [6] H.Rusinck, A. Levy, and M.E.No. "Performance of two methods for registering PET and MR brain scans", *Conf.Record 1991 IEEE Nuclear Science Symp. Med. Imag. Conf.*, vol. 3, pp. 2159-2162, 1991
- [7] A.V.Oppenheim and J.S.Lim, "The importance of phase in signals", *IEEE Proc.*, vol. 69, no. 5, pp 529-541, 1981
- [8] B.V.K.V. Kumar and E. Pochapsky, "Signal-to-noise ratio considerations in modified matched spatial filters", *J.Opt.Soc. Am. A.*, vol. 3, no.6, pp. 777-786, 1986
- [9] B.S.Reddy, B.N.Chatterji, "An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration", *IEEE Trans. On Image Processing*, vol.5, no.8 pp. 1266-1271, 1996