움직이는 객체를 포함하는 영상의 컨투어 기반 모자이킹 방법

정 성 룡, 최 윤 희, 최 태 선
광주과학기술원 기전공학과
전화 : 062-970-2419 / 핸드폰 : 016-620-3728

Contour-Based approach for mosaicking images that contain moving objects

Sung Yong Jung, Yun-Hee Choi, Tae-Sun Choi
Dept. of Mechatronics, Kwang-Ju Institute of Science and Technology (K-JIST)
E-mail : jaysr@naver.com

Abstract

This paper has been studied how to deal with moving objects in images when we mosaic them. The global motion between two images is biased due to the local motion from these moving objects, so it is very important how to eliminate the effects of them. In this paper contour-based approach for mosaicking images that contains moving objects is presented. Once we get the contours of images we can both eliminate the moving objects and mosaic the images. In this stage, hierarchical moving objects elimination technique is introduced. Experiment is done for Stefan tennis sequences to verify the proposed algorithm.

I. 서론

움직임 추정 기법은 영상 데이터의 효율적인 저장과 전송을 위해 오랫동안 심도있게 연구되어 왔다. 움직임에는 로컬 모션과 글로벌 모션이 있으며, 이미지 모자이킹(또는 이미지 정합)은 글로벌 모션의 추정을 그 목적으로 한다. 테니스 경기 중계방송과 같이 운동장이라는 동일 배경영역에 대한 시간적 중복성이 상당한 경우, 이미지 모자이킹을 통해 전체배경을 한번만 구성해 놓으면 그 다음부터 각 프레임은 그 전체배경에 대한 모션벡터 정보만 가지면 되므로 매우 효율적인 데이터의 저장과 전송을 꾀할 수 있다. 또한 글로벌 모션이 갑자기 변화하는 경우를 토대로 장면전환을 검출할 수도 있다.

본 논문에서는 움직이는 객체를 포함한 동영상 프레임들 간의 모자이킹을 각 프레임의 컨투어 정보를 이용하여 실현하였다. 영상들이 움직이는 객체들을 포함하는 경우, 그 객체들에 따른 로컬 모션으로 인해 전체 영상간의 글로벌 모션이 영향을 받는다 [1,3]. 본 논문에서는 영상의 컨투어 정보를 이용하여 로컬 모션을 효과적으로 제거하고 글로벌 모션만을 추정해내는 것을 실현하였다.

II. 서로 대응하는 컨투어 쌍 찾기

이번 절에서는 모자이킹 하고자 하는 두 영상간의 대응관계(또는 매칭관계)를 찾는 과정이 기술된다. 이전 영상의 어떤 부분이 현재 영상의 어떤 부분과 대응하는지를 알아내야, 두 영상 간의 글로벌 모션을 계산할 수 있다. 본 논문에서는 컨투어 정보를 이용하여 두 영상간의 대응관계를 알아내었다.

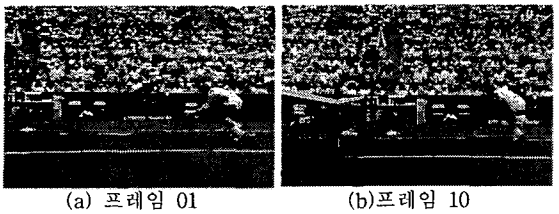
2.1 컨투어 추출

각 영상의 컨투어를 추출하기 위하여 LoG(Laplacian of Gaussian) 연산자가 사용되었다.

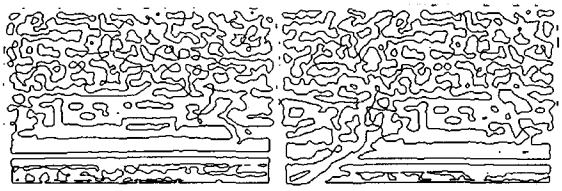
$$LoG = \left(\frac{r^2 - \sigma^2}{\sigma^4} \right) \exp\left(-\frac{r^2}{2\sigma^2} \right) \qquad (2.1)$$

여기서 $r^2 = x^2 + y^2$ 이고 σ는 표준편차이다.

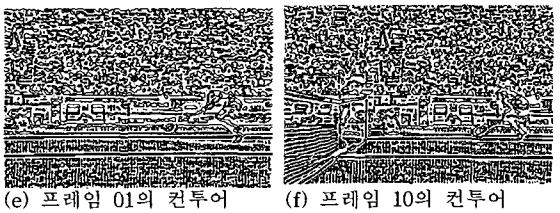
이 연산자의 사용에 있어서는, 가우시안 연산자의 σ값과 라플라시안 연산자에 대한 문턱값 ThL을 어떻게 설정하는지가 매우 중요하다. 이에 따라 추출되는 컨투어가 판이하게 달라지기 때문이다.



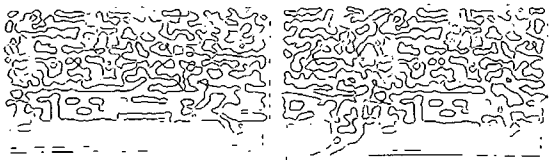
(a) 프레임 01 (b)프레임 10



(c) 프레임 01의 컨투어 영상 (σ=4, ThL=0) (d) 프레임 10의 컨투어 영상 (σ=4, ThL=0)



(e) 프레임 01의 컨투어 영상 (σ=1, ThL=0) (f) 프레임 10의 컨투어 영상 (σ=1, ThL=0)



(g) 프레임 01의 컨투어 영상 (σ=4, ThL=0.0006) (h) 프레임 10의 컨투어 영상 (σ=4, ThL=0.0006)

그림 2.1. 스테판 테스트 시퀀스에 대한 컨투어 영상

그림 2.1(e),(f)에서 보는 바와 같이 너무 작은 σ값은 대응관계를 찾는 데 별로 도움이 안 되는 사소한 컨투어를 너무 많이 발생시키며, 그림 2.1(c),(d)와 (g),(h)를 비교하면 알 수 있듯이 ThL의 값을 0으로 설정하는 경우에만 닫힌 컨투어들로만 구성된 컨투어 영상을 얻을 수 있다.

논문 [2]에서는 좀 더 정확한 컨투어 매칭을 위해 좀 더 많은 수의 컨투어를 발생시키고자 비교적 낮은 σ 값인 3을 사용하였다. 따라서 사소한 컨투어들이 너무 많이 발생하게 되어 ThL을 0이 아닌 값으로 설정하다 보니, 그림2.1(g),(h)와 같이 열린 컨투어들까지 포함하는 컨투어 영상을 얻게 되었다. 이에 따라 매칭 컨투어 쌍을 알아내기 위해, 닫힌 컨투어에 대한 대응관계를 우선 찾아내고 이를 초기값으로 하여 열린 컨투어 간의 대응관계를 찾아내는 등 알고리즘이 복잡하게 되었다.

본 논문에서는 σ값을 보다 높은 값인 4로 설정함으로써 ThL을 0으로 설정해도 사소한 컨투어들이 거의 발생하지 않게 하였으며, 이에 따라 컨투어 매칭은 닫힌 컨투어의 질량중심에 대해서만 수행되면 되므로 알고리즘이 많이 간소화되었다. 이 방법은 [2]와 비교할 때 컨투어(닫힌 것이든, 열린 것이든) 자체의 개수를 감소시켜 잘못된 컨투어 매칭에 대한 민감도가 좀 더 높아지기는 하나, 5를 넘지 않도록 σ값을 설정하고 앞으로 설명할 일관성 체크 방법을 적용하면 오류가 없는 결과 모자이크 영상(그림3.1)을 얻을 수 있음을 확인하였다.

2.2. 매칭 컨투어 쌍 구하기

이제 우리의 목표는 그림 2.1(c),(d)와 같이 생성된 컨투어들 사이의 대응관계를 파악하여, 그 대응관계로부터 두 영상 간에 확대축소, 회전, 병진, 비틀림 등이 얼마만큼 일어났는지를 알아내는 것이다. 이것은 원근변환으로써 표현되며, 궁극적으로 우리는 원근변환 모션 파라미터인 $a_1 \sim a_8$를 구해내야 한다. 원근변환에 대한 식은 다음과 같다.

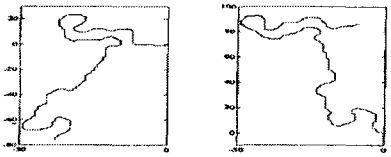
$$x' = \frac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + 1}$$

$$y' = \frac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + 1} \qquad (2.2)$$

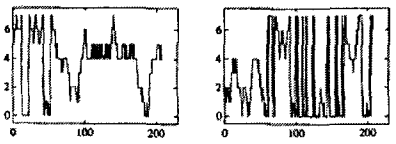
여기서 (x',y')와 (x,y)는 컨투어의 질량중심이다.

컨투어들 간의 대응관계를 알아내기 위해서 본 논문에서는 정규화된 체인코드를 사용하였다. 그림 2.2에서 보는 바와 같이 정규화된 체인코드는 두 영상 간에 회전변환이 일어났을 경우에도 적용할 수 있는 장점이 있다 [2].

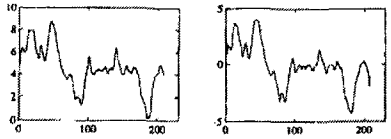
비교하고자 하는 두 개의 컨투어가 서로 다른 개수의 픽셀로 이루어졌다면, 보간법을 이용하여 작은 개수 쪽으로 서로 개수를 맞춘 후 정규화된 체인코드 간



(a) 대응되는 컨투어 쌍



(b) (a)에 대한 체인코드



(c) (a)에 대한 정규화된 체인코드

그림 2.2. 정규화된 체인코드

의 유사성 비교를 한다. 이 때 유사성 정도가 일정 문턱값을 넘어설 경우, 매칭 컨투어 쌍 후보로 삼는다 [2].

III. 계층적 일관성 검사를 통한 움직이는 객체의 제거

이제 2.2절에서 얻은 매칭 컨투어 쌍 중에서 움직이는 객체로부터 기인한 매칭 컨투어 쌍을 제거하는 방법에 대해 기술한다. 움직이지 않는 배경으로부터 기인한 모션들은 모두 동일할 것이고, 움직이는 객체로부터 기인한 모션만이 현저히 다른 모션을 가질 것이므로 이를 이용해 움직이는 객체의 모션을 제거한다.

모션의 일관성에 근거한 방법이므로, 일관성 검사라 칭한다. 본 논문에서는 기존에 제시된 강체변환을 통한 일관성 검사방법 [1]에 추가적으로, 컨투어의 구성 픽셀 개수에 따른 일관성 검사방법을 도입하여 움직이는 객체의 모션을 더욱 정확히 제거하도록 하였다.

3.1. 컨투어의 구성 픽셀 개수에 따른 일관성 검사

정규화된 체인코드 간의 유사성 검사를 마친 매칭 컨투어 쌍들은 일관된 확대축소율을 가질 것이다. 즉, 컨투어를 구성하는 픽셀의 개수 비율이 각 매칭 컨투어 쌍별로 일관된 값을 가질 것이다. 만약 다른 확대축소율을 보이는 컨투어 쌍이 있다면, 그것은 움직이는 객체로부터 기인했다고 판단할 수 있다.

제안한 일관성 검사를 위해서는 우선 각 컨투어 쌍에 대한 픽셀개수의 비율, 즉 확대축소율 *PNRatio*을 계산한다.

$$PNRatio(i) = PN1(i) / PN2(i)$$
$$\text{for } i = 1,2,\ldots,N \qquad (3.1)$$

여기서 N은 매칭 컨투어 쌍들의 총 개수이고, *PN1(i)*, *PN2(i)*는 각각 이전 이미지에서 i번째 컨투어 쌍을 구성하는 픽셀의 개수, 현재 이미지에서 i번째 컨투어 쌍을 구성하는 픽셀의 개수이다. 결과로서 *PNRatio*는 1xN 행렬이 나온다. 이제 이 확대축소율의 대표값을 선정하기 위해 각 컨투어의 픽셀 개수 비율 *PNRatio(i)*로 *PNRatio* 전체를 나눈다. 수식은 다음과 같다. 결과로는 NxN 행렬이 나온다.

$$PNRatio5(i,j) = \{PNRatio(1,j) / PNRatio(1,i)\}^5$$
$$\text{for } i = 1,2,\ldots,N,\ j = 1,2,\ldots,N \qquad (3.2)$$

이 때 확대축소율의 차이를 확연히 구분하기 위해서 5제곱을 하여 *PNRatio5*를 얻었다.

이제 *j=1,2,...,N*에 대해 다음과 같은 수식을 만족시키는 회수가 가장 많은 *PNRatio(i)*를 확대축소율의 대표값 *PNRRepre*로 취한다.

$$\frac{1}{Th1D} \le PNRatio5(i,j) \le Th1D$$
$$\text{for } j = 1,2,\ldots,N \qquad (3.3)$$

이 때, *Th1D*는 1.2로 놓는 것이 가장 적절하였다.

$$PNRRepre = PNRatio(i) \qquad (3.4)$$

참고로, *PNRRepre*는 1x1 행렬이다. 이제 확대축소율의 대표값을 찾았으면 그 값으로 전체 *PNRatio*을 나눈 후 다음 기준에서 벗어나는 컨투어 쌍을 움직이는 객체에서 기인한 것이라 보고 제거하면 된다.

$$\frac{PNRatio(i)}{PNRRepre} \ge C1D \quad \text{or} \quad \frac{PNRatio(i)}{PNRRepre} \le \frac{1}{C1D}$$
$$\text{for } i = 1,2,\ldots,N \qquad (3.5)$$

여기서 *C1D*는 적절한 상수값이다.

이제 몇 쌍이 제거된 매칭 컨투어 쌍들을 가지고 다시 식(3.1)부터 식(3.5)까지 반복한다. 식(3.5)를 적용함에 있어 제거할 컨투어 쌍이 더 이상 생기지 않을 경우, 3.2절에서 소개할 강체변환을 이용한 일관성 검사 단계로 넘어간다.

3.2. 강체변환을 이용한 일관성 검사

컨투어를 구성하는 픽셀 개수에 의한 일관성 검사는 단 하나의 컨투어 쌍만 있어도 검사척도를 마련할 수 있으므로 1차원적인 검사라 할 수 있고, 강체변환을 이용한 일관성 검사는 최소 2쌍의 컨투어 쌍이 있어야만 검사척도를 마련할 수 있으므로 2차원적인 검사라 할 수 있다. 이것은 강체변환 모델이 다음과 같이 표현되기 때문이다.

$$x' = a_1 x - a_2 y + a_3$$
$$y' = a_2 x + a_1 y + a_4 \qquad (3.6)$$

실험에 따르면, 낮은 차원의 검사인 3.1절의 검사 후에 높은 차원의 검사인 3.2.절의 검사를 수행할 경우 움직이는 객체의 제거가 보다 오류 없이 수행되었다. 그러나 반대의 순서로는 오류율이 감소하지 않았다. 즉 낮은 차원에서부터 높은 차원으로의 계층적 일관성 검사는 움직이는 객체의 제거에 대한 오류율의 측면에 있어서 신뢰성을 높여줌을 확인하였다.

강체변환을 통한 일관성 검사 [1]는 다음과 같이 수행한다. 우선 3.1절의 검사를 거친 매칭 컨투어 쌍 후보들의 무게중심 (x_i,y_i)와 (x_i',y_i')들과 식(3.6)으로부터, 최소제곱 방법을 이용해 강체변환 파라미터의 초기값을 얻는다. 구해진 모션 파라미터를 이용해서 (x_i,y_i)에 대한 예측되는 대응좌표 (x_i'',y_i'')를 얻는다. 이제 예측되는 대응좌표 (x_i'',y_i'')와 실제의 대응좌표 (x_i',y_i') 사이의 차이가 현저하다면, 이것은 움직이는 객체로부터 기인한 매칭 컨투어 쌍이라 판단하고 제거하면 된다. 검사척도 d_i는 다음과 같다.

$$d_i = \sqrt{(x_i'' - x_i')^2 + (y_i'' - y_i')^2} \qquad (3.7)$$

이제 d_i을 문턱값 *Th2D*와 비교한다. 만약 d_i이 *Th2D*보다 크다면 그에 해당하는 컨투어 쌍을 제거하면 되는 것이다. 다음은 *Th2D*에 대한 공식이다.

$$Th2D = \frac{1}{N}\sum_{i=1}^{N} d_i + C2D\sigma_d \qquad (3.8)$$

여기서 σ_d는 d_i에 대한 표준편차이고, *C2D*는 적절한 상수값이다.

이제 남은 컨투어 쌍들에 대해 다시 강체변환 파라미터를 구한 다음, 식(3.7)과 식(3.8)을 제거되는 컨투어 쌍이 없을 때까지 반복한다.

3.3. 실험결과

그림 3.1은 테니스 경기의 프레임 01 과 프레임 10에 대해 제안된 알고리즘을 사용하여 모자이킹을 수행한 결과이다.

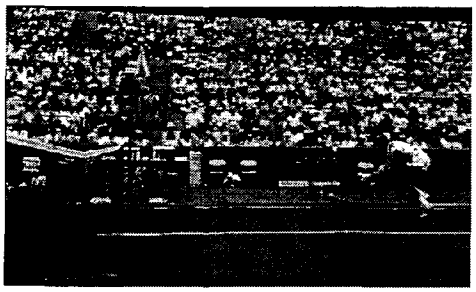


그림 3.1. 프레임 01 (그림 2.1(a)) 와 프레임 10 (그림 2.1(b)) 의 모자이크 영상

Ⅳ. 결론

기존의 이미지 모자이킹 방법들은 주로 움직이는 객체를 포함하지 않은 영상에 대한 제안들이었다. 그러나 영상이 움직이는 객체를 포함하게 되면, 이로 인한 로컬 모션의 발생으로 인해 글로벌 모션이 편향되게 된다. 따라서 이러한 로컬 모션을 효율적으로 잘 제거하는 것이 전체 글로벌 모션 추정의 정확도에 결정적 역할을 한다.

본 논문에서는 컨투어를 구함으로써, 이를 움직이는 객체의 제거에도 이용하고 모자이킹에도 이용하였다. 또한 계층적 일관성 검사방법을 도입해, 매칭 컨투어를 더욱 정확히 찾을 수 있도록 하였다. 컨투어 방법은 서로 다른 센서를 통해 얻은 영상들 또는 영상전체에 밝기 변화가 일어난 영상들 간의 모자이킹까지 가능한 접근방법이므로, 위성영상처리 등의 특별한 응용분야에는 더욱 강력하게 사용될 수 있을 것이다.

참고문헌

[1] Chiou-Ting Hsu, "Mosaics of video sequences with moving objects", IEEE International Conference on Image Processing, Vol. 2, pp. 387-390, 2001

[2] Hui Li, "A contour-based approach to multisensor image registration", IEEE Transactions on Image Processing, Vol. 4, pp.320-334, March 1995

[3] Davis, J., "Mosaics of scenes with moving objects", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings, pp.354-360, 1998

[4] Irani, M.,; Anandan, P.; Hsu, S., "Mosaic based representations of video sequences and their applications", IEEE International Conference on Computer Vision, Fifth Proceedings, pp. 605-611, 1995

[5] Szeliski, R., "Video mosaics for virtual environments", IEEE Computer Graphics and Applications , Vol. 16, pp. 22-30, March 1996

[6] Hyunwoo Ki; Ki Sang Hong, "Soccer video mosaicing using self-calibration and line tracking", IEEE International Conference on Pattern Recognition, 15th Proceedings, Vol. 1, pp. 592-595, 2000