

모폴로지 코너 검출법을 이용한 영상 모자이크

조세연[○], 이정호, 유형승, 조아영, 정동석
인하대학교 전자공학과

sefungs@hanmail.net[○], gglee@nownuri.net, helf-me@hanmail.net, aqua0654@hotmail.com, dsjeong@inha.ac.kr

Image Mosaics using Morphological Corner Detection

Se-yeon Jo[○], Jung-ho Lee, Hyung-sung Yu, A-yeong Jo, Dong-seok Jeong
Dept. of Electronic Engineering, Inha University

요 약

모자이크는 여러 장의 영상을 하나의 큰 영상으로 만드는 것을 말한다. 본 논문은 asymmetrical closing이라고 불리는 모폴로지에 의한 closing operator를 사용한 영상 모자이크에 관한 연구이다. asymmetrical closing을 하기 위한 structuring element를 소개하고 이것을 이용한 코너 점 추출 방법 및 local maxima에 대해서도 소개 한다. 여러 개의 코너 점들 중 조건을 만족하는 tie point들을 이용하여 perspective 변환 파라미터를 추출하여 최종 모자이크 결과 영상을 생성하게 된다.

1. 서 론

영상 모자이크 기술은 컴퓨터 그래픽스, 컴퓨터 비전, 멀티미디어 시스템의 응용분야중 하나이다. 카메라의 시야는 사람의 눈에 비해 좁은 범위를 가지고 있으므로, 여러 장의 영상을 연결하면 한 장의 영상으로는 획득하지 못하는 몇몇 큰 물체들을 한 장의 큰 영상 안에 담을 수가 있다. 본 논문에서 제안하는 방법은 다음과 같이 구성된다. 첫째로, image pyramid 방법을 이용한 여러 단계의 매칭 지점을 이용하여 탐색영역을 줄여 효과적으로 탐색하였다. 두 번째로 모폴로지 코너 검출방법을 이용하여 붙이려는 두 영상의 후보 tie-point들을 추출 하였다. 세 번째로 각 영상에 대해 가장 넓은 범위의 4개의 tie point들을 검출하는 ABM(Area Based Matching) 방법을 적용한다. 마지막으로 transformation parameter를 추출하여 세 개의 입력영상을 모자이크 한다.

2. 본 론

2.1 Image Correlation

Image Correlation은 그레이 영상에서 MSE(Mean Square Error) 방법을 사용하였다. 영상 간 중첩 거리를 1에서 점점 늘려가며 MSE를 계산하여 그 값이 최소가 되는 중첩거리를 결정하였다. 아래의 (1)식은 MSE 계산 식이다. 이때, 전체 영상에 대하여 MSE를 계산할 경우, 연산량이 많기 때문에 계산 속도를 향상시키기 위하여 image pyramid 방법을 사용하였으며, 그레이 영상의

$$\sigma_{MS}^2 = E[|u(m, n) - u'(m, n)|^2]$$

σ_{MS} : Mean Square Error.

(1)

E : Expectation.

u : Reference Image, u' : Matched Image.

1/4, 1/16 크기로 pyramid image를 생성한다. 작은 크기의 pyramid image로부터 중첩 거리를 계산하여 상위 단계로 올라가며 보다 정확한 중첩 거리를 계산한다. 상위의 단계에서 매칭 지점을 이용하여 다음 영역의 탐색 영역을 줄여 효과적으로 탐색하였다.

2.2 Corner Detection -Asymmetrical Closing

Perspective transformation을 하려면 영상이 겹쳐지는 각 부분에서 대응되는 tie-point 4개가 필요하다. 여기에서 제시하는 방법은 morphological structure를 이용한 asymmetrical closing을 이용해 코너 점들을 검출하는 기법이다. 전역 문턱치를 이용할 경우 코너 점이 과도하게 검출되는 문제점을 해결 하기위해 local maxima방법을 제안한다. asymmetrical closing이란 그림 1의 structuring element를 중 서로 다른 두 가지 element를 이용하여 dilation 과 erosion을 순차적으로 수행하는 것을 말하는데 dilation을 할 때 structuring element를 이용하여 masking operation을 하면서 마스크가 씌워진 화소의 값들 중 가장 큰 값을 중앙 화소의 값으로 지정하며 $I^o(x, y)$ 으로 나타낸다. 또 erosion을 통해 마스크가 씌워진 화소의 값들 중 가장 작은 값을 중앙화소의 값으로 지정하며 $I^e(x, y)$ 으로 나타낸다. 그림1의 (a) 및

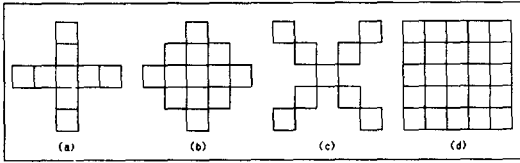


그림 1. (a) structuring element, +. (b) ◇. (c) X. (d) □.

$$I_{+, \diamond}^c = (I_+^{\delta})_{\diamond}^c \quad (2)$$

$$I_{X, \square}^c = (I_X^{\delta})_{\square}^c \quad (3)$$

(b)를 이용한 closing은, 식(2)와 같으며, (c) 및 (d)를 이용한 closing은, 식(3)과 같다. 그 결과를 보여주는 실험 영상은 그림 2의 (a), (b)와 같다. (c)는 원본 영상이며 (d)는 특정 문턱치 이상의 corner strength를 가진 화소들을 나타낸다. corner strength란 어떤 화소에 대한 서로 다른 두 closing의 차이 값을 나타낸다[1].

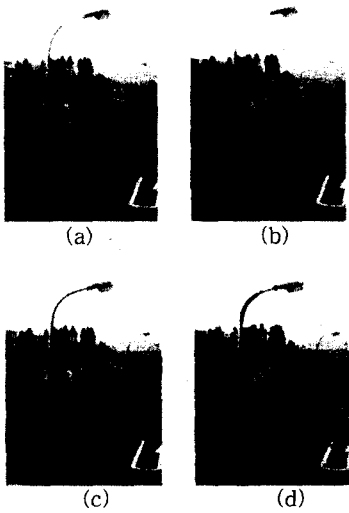


그림2. (a) +, ◇ element를 이용한 closing. (b) X, □ 이용한 closing. (c) 원본영상. (d) corner strength가 20 이상인 화소를 표시한 영상.

2.3 Corner Detection - Local Maxima

Corner detection을 할때 전역 문턱치를 이용할 경우 문턱치 값에 따라 검출되는 코너 점 개수가 달라진다는 문제점이 있다. 그래서 local maxima를 이용한 코너 점 검출을 시도했다. 아래 그림3은 Corner Strength 분포이다. 아래에 보이는 것처럼 코너 점을 중심으로 산처럼

corner strength 값이 분포해 있다. 각 산의 봉우리 지점(local maxima)을 검출하여 해당 지점을 그 지역 영역의 코너 점으로 정한다.

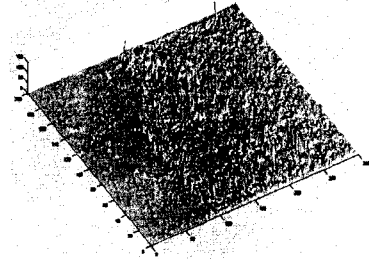


그림 3. 예제 영상의 Corner Strength 분포

2.4 Corner Detection - Tie Point 검출

다음은 local maxima를 이용한 후보 tie point 들 중에 왼쪽 영상의 화소와 일치하는 화소를 붙이려는 영상에서 찾는 작업이 필요하다. 왼쪽 영상의 화소와 매칭되는 화소를 오른쪽 영상에서 찾기 위해 미리 정의한 search window(왼쪽 영상의 corner point로 구한 특징점들을 중심으로 60*60의 search window)를 이동시키면서 오른쪽 영상의 patch 인 template 영상(오른쪽 영상 코너 점의 특징점들을 중심으로 60*60의 template 영상)에 대응하는 영상을 찾는다. 왼쪽 영상의 search window와 오른쪽 영상의 template 영상이 중첩 구간을 벗어나지 않게 하기 위해 중첩 구간의 바깥쪽으로부터 30화소 구간 안쪽의 특징점들만 고려한다. search window를 이동시키면서 template 속에서 정의된 화소를 결정하기 위해서 농담정규화 정합법(Normalized Gray-level Correlation: NGC)을 사용하였다. NGC란 입력 영상에서 취한 비교할 영상 영역의 밝기 평균을 구해서 평균값을 빼버린 영상을 만들고 이 영상을 비교하기 위한 영상으로 사용하는 방법으로써 식(4)와 같다.

$$r = \frac{a \cdot b}{|a| \cdot |b|} = \frac{\sum_{i=0}^M \sum_{j=0}^N a(i, j)b(i, j)}{\sqrt{\sum_{i=0}^M \sum_{j=0}^N a(i, j)^2 \sum_{i=0}^M \sum_{j=0}^N b(i, j)^2}} \quad (4)$$

$$a(i, j) = g(i, j) - \bar{m}, \quad b(i, j) = t(i, j) - \bar{i}$$

$a(i, j)$ 는 입력 영상에서 취한 비교부에서 평균 \bar{m} 을 뺀 밝기이며, $b(i, j)$ 는 template 영상에서 평균 밝기 \bar{i} 를 뺀 값을 나타낸다. 즉, 평균 밝기를 빼서 정규화시킨 두 영상의 두 벡터 비교식을 이용하여 유사도를 검사한다.

2.5 Image Warping

Image warping은 서로 다른 여러 장면들을 하나의 큰 장면으로 만들고, 블렌딩 하는 작업을 의미한다. 먼저 합치려는 영상간의 관계를 표현 할 수 있는 geometric 변환을 고려해야 한다. image plane 상의 2D 점들을 (x, y, w)라고 표시하고, 거기에 상응하는 cartesian 좌표계를 (x/w, y/w)라고 한다[2]. 아울러 homogeneous 좌표계를 사용하여, 2D planar projective 변환을 행렬의 곱을 사용하여 식 (5)로 표현 할 수 있다.

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{pmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \quad (5)$$

위 식의 m_0 에서 m_7 까지의 8개의 파라미터를 결정하기 위한 첫 번째 작업으로 모폴로지 코너 검출법을 이용하여 추출한 Tie Point들을 이용하여 파라미터들을 결정하는 것이다. 우리가 붙이려는 두 장의 영상을 각각 A, B라고 가장 적합한 Tie Point를 4개씩 가지고 있다면 8개의 선형방정식을 생각할 수 있고 8개의 파라미터들을 결정할 수가 있다[3].

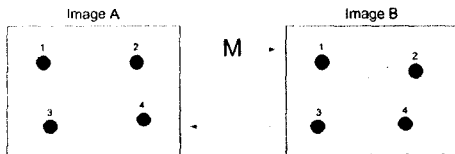


그림 4. 영상간의 관계 나타내는 행렬 M

그림 4에서 영상 A를 참조 영상, 영상 B를 붙이려는 영상이라고 하고 그들 사이의 관계를 나타내는 변환 행렬을 M이라 하면, 크기는 3*3 정방행렬이다[4]. 일단 파라미터를 결정 하기위해서 일단 $X = Hx$ 라는 식을 생각해 보자. X는 world plane 좌표를 의미하고, x는 image plane 좌표를 의미한다. H는 행렬변환을 나타낸다. $X = Hx$ 라는 식을 더 확장해 보면 식(6)과 같이 나타낼 수가 있다[5].

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -X_3x_3 & -X_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -Y_3x_3 & -Y_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -X_4x_4 & -X_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -Y_4x_4 & -Y_4y_4 \end{pmatrix} \begin{pmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \\ m_7 \end{pmatrix} = \begin{pmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ X_3 \\ Y_3 \\ X_4 \\ Y_4 \end{pmatrix} \quad (6)$$

영상 블렌딩 은 두 영상을 합칠 때 생성되는 영상의 화소 값을 결정하는 방법으로서, 렌즈의 왜곡으로 발생하는 고스트를 제거한다. 여기서 우리는 중첩 거리에 따라

선형적으로 보간 되는 선형가중치 함수를 사용하였다.

3. 결과 영상

그림 5는 서로 다른 곳에 위치한 세 개의 카메라로부터 받아들인 (a),(b),(c) 입력 영상을 모자이크 처리한 결과를 (d)에서 보여주고 있다. (d) 영상을 보면 왼쪽 영상과 중앙 영상 간 파노라마 영상에 약간의 고스트가 발생한 것을 볼수

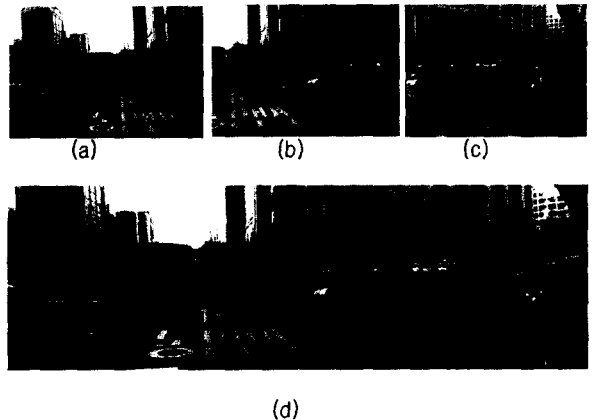


그림 5. (a) 왼쪽 입력영상 (b) 중앙 영상 (c) 오른쪽 영상 (d) 결과 모자이크 영상

가 있다. 하지만 대체적으로 만족 할만한 파노라마 영상을 얻었다.

4. 결론

모폴로지를 이용하여 영상의 코너 점들을 추출한 후 Template 매칭을 이용하여 각 영상마다 4개씩 특징 점들을 추출하였다. 연산량이 많아 수행시간이 길었다. 수행 시간을 줄이기 위한 고찰이 필요 할 것 같다. 그리고 4개의 특징 점들을 가지고 8개의 파라미터들을 결정하기 위해 각 영상에 대해 가장 넓은 범위의 tie point들을 검출하여야 한다. 만약에 4개의 tie point가 작은 영역에 걸쳐 생성 되었을 때 행렬 H가 잘못 생성되어 좋지 않은 결과 영상을 얻게 되므로 주의 해야 한다.

참고문헌

- [1] R.Laganieri, "A Morphological operator for Corner Detection," Pattern Recognition, vol.32, No.11, pp. 1643-1652, 1998.
- [2] J.D.Foley et al, "Computer Graphics:Principles and Practice," Addison-Wesley, 1990.
- [3] R.Szeliski, "Video Mosaics for Virtual Environments ," IEEE CG&A, pp.22-30, March 1996.
- [4] B.Itkowitz, "Image Mosaics," CAS CS585 Assignment P4, December 1998.
- [5] <http://xenia.media.mit.edu/~cwren/interpolator>