

Hausdorff Distance를 이용한 특징기반 영상 모자이킹

고 종호, 이 철우
전남대학교 컴퓨터공학과

Feature-based Image Mosaicing Using Hausdorff Distance

Jong-Ho Ko, Chil-Woo Lee

* Dept. of Computer Engineering, Chonnam National Univ.
{jhko, leecw}@image.chonnam.ac.kr

요 약

본 논문은 움직임이나 장면이 제약되지 않는 카메라로부터 획득한 여러장의 디지털 영상을 한장의 통합된 영상으로 자동적으로 모자이킹 하는 방법에 대해 기술한다. 우선 영상을 세그멘테이션과 레이블링으로 영상의 영역을 분할하고, 특징 점을 찾는다. 같은 영역으로 분할된 곳의 특징점을 각각의 그룹으로 하고 영역안의 특징점을 hausdorff distance를 이용 영상을 비교하여 대응점을 구한 후 비선형 이송오차 최적화 알고리즘을 이용해 최적의 변환 행렬로 하나의 영상을 자동적으로 구성하는 과정에 대해서 기술한다.

1. 서론

최근 몇 년 전부터 영상 모자이킹 기술은 컴퓨터 비전 연구에서 많이 응용이 되기 때문에 흥미 있는 연구 분야가 되어 왔다. 이 기술은 우주/항공 사진의 해석, 컴퓨터 그래픽스의 페인팅 시스템, 사진의 가공 등 영상을 다루는 분야에 폭 넓게 사용되어 왔다.

모자이킹 기술은 임의의 시점에서 촬영한 영상들을 하나의 영상으로 통합시키는 물론 시점을 달리하여 영상을 표현할 수 있기 때문에 건축물 내부와 같이 복잡한 장면을 그래픽 모델을 제작하지 않고도 간단한 변환

에 의해 생성할 수 있으므로 다양한 영상 생성이 필요한 각종 응용 시스템에서 사용될 수 있다. 이러한 기술을 최근에 들어 영상기반 렌더링(image based rendering) 이라고 부르며 계산의 복잡도를 줄이는 문제와 3차원 정보를 포함하는 사실성(reality) 중대가 당면 과제로 남아 있고, 지속적인 연구가 계속되고 있다.

지금까지 널리 알려진 모자이킹 방법으로는 파노라마 영상 생성법이 있다[6]. 이 방법은 고정된 위치에 카메라를 설치하고 수평방향으로 회전시키면서 촬영한 영상을 하나의 영상으로 통합하여 원통형의 벽에 매핑 시킴으로써 관찰자에게 전 방위의 시야를 제공하는 방법이다. 하지만, 이 방법은 카메라가 자유롭게 못하고, 영상 획득시 고려해야 할 제한된 조건이 많이 존재한다.

여러가지의 영상을 한장의 모자이크 영상으로 생성을 위해 초기에는 사용자가 수작업으로 각각의 영상 쌍마다 대응점을 지정하였으나 완전한 자동화를 실현하기 위해 많은 연구가 진행되고 있고, 더 나아가 3차원의 현실감 있는 영상을 제공하기 위해 영상 생성시의 카메라 파라미터를 복원하는 연구도 진행되고 있다.

모자이킹 기술은 근본적으로 그래픽스에서 사용하는 영상 워핑(Image Warping) 이론과 컴퓨터 비전에서 사용하는 투영변환 (Perspective Transformation) 이론을 결합한 기술이다. 이 기술에서는 임의의 영상을 새로운 영상으로 변환하는 과정에서 각종 파라미터를 고려해야 되나 파라미터가 많으면 많을수록 계산이 복잡하여 처리시간이 길어질 뿐 만 아니라, 국소 극소점에 수

림할 가능성이 있어 바람직하지 않다. 이러한 문제점을 해결하기 위해, Szeliski[1]는 두 영상 사이의 회전만을 고려한 3-파라미터 회전 모델을 제시하였다. 이 모델은 두 영상의 초점 거리를 임의로 가정하고, 그 값은 별도의 계산이 필요한 최소화 방법[7]에 의해 초기 추정치로 부터 반복적으로 개선된다. Chiba[5]는 다중 광류(optical flow) 추정법을 이용하여 모자이크하는 방법을 소개하였다. 먼저, 초기 중첩영역을 구하고, 광류추정의 한 방법인 Lucas-Kanade 방법을 이용 대응점을 찾는다. 이로부터 변환행렬을 구해 에피폴라 기하에 적용시켜 영상을 모자이크하는 방법을 소개하였다.

본 논문에서는 기존에 제시되었던 영상 모자이크 기술과는 달리 자유로운 카메라로부터 획득한 두 장의 영상을 세그멘테이션을 통한 영역을 분리하고, 분리된 영역에서의 특징점을 쌍으로 하여 Hausdorff distance를 계산한 후 영역간의 매칭을 통한 영상모자이크의 전체 과정을 인간의 개입없이 자동적으로 모자의 영상을 생성하는 방법에 대해 소개한다. 먼저, 통합할 두장의 영상을 각각 세그멘테이션과 특징 점을 찾는다. 다음으로, 영역안에서의 추출된 특징 점들 간의 관계를 Hausdorff distance를 통해 최대의 유사도를 측정, 영역에대한 4개의 대응쌍을 찾고, 4개의 대응점들을 투영변환을 통해서 비선형 이송오차 최소화 알고리즘을 이용, 최적의 변환 행렬을 구한 후 최종적으로 영상을 합성한다.

본 논문의 구성은 다음과 같다. 2절에서는 영상 모자이크를 수행하기 위한 전체적인 과정에 대해서 기술한다. 3절에서는 실험 결과와 고찰에 대해서 서술하고, 4절에서는 결론을 맺는다.

2. 영역기반 영상모자이크

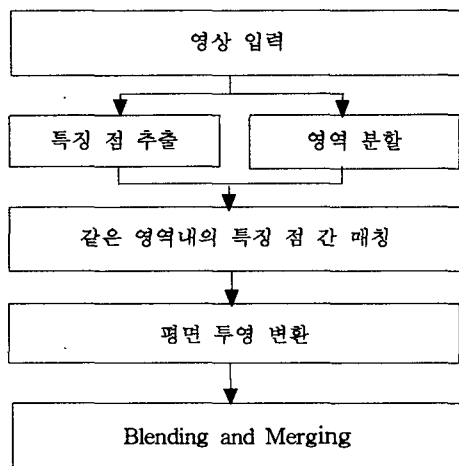


그림 1. 자동 모자의 영상 생성 과정

그림1. 은 본 논문에서 제시한 자동 영상 모자이크 생성의 전체적인 흐름이다.

2.1 특징점 추출(Feature Point Extraction)

특징 점이란 영상 안의 물체가 가지고 있는 코너 점(Corner Point)이나 주변 화소에 비해 큰 밝기의 변화를 갖는 화소와 같은 점으로써 이 점은 영상 안에서 다른 점들과 구별되어지는 점이다. 이러한 특징 점은 여러 방향에서의 높은 미분 값을 갖는 점으로 정의되기도 하고, 특정 변환에 불변의 성질을 갖는 점으로 정의되기도 하는 등 여러 가지로 정의될 수 있다. 영상 안에서 특징 점을 뽑아내는 방법들은 다음과 같이 크게 세 가지 그룹으로 분류할 수 있다.

- 그룹 1 : 에지 지도(Edge Map)를 이용한 방법
- 그룹 2 : 그레이레벨(Gray Level) 영상을 직접 접근하는 방법
- 그룹 3 : 미분 기하 또는 모델 기반 접근법

이 가운데 대부분의 특징 점 추출은 두 번째 그룹에 속하며, 최근에는 그룹 3에서 활발한 연구가 진행되고 있다. 그룹 3에서 미분기하 방법은 Guiducci 방법으로 그룹 2에서 파생된 방법이고, 모델 기반 법은 Gaussian filter를 이용 smooth를 수행 후 분석하는 방법으로써 잘 알려진 정확한 코너를 찾는데 이용하지만, 계산시간이 많이 걸린다는 단점을 가지고 있다.[4]

본 논문에서는 그룹 2에 속한 방법으로 Hessian Matrix 이용 방법을 제안한 Beaudet 코너 검출기를 이용했다. 이 방법은 영상의 밝기가 급격히 변화는 점(코너 점)을 특징 점 후보로 선택한다. Hessian Matrix는 식 (1)과 같다.

$$H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix} \quad \text{식 (1)}$$

여기에서, I_{xx} 와 I_{yy} 는 각각 함수 I 의 x 방향과 y 방향의 2차 미분이며, I_{xy} 와 I_{yx} 는 함수 I 의 x 방향과 y 방향의 2차 미분이다. 헷세 행렬의 행렬식 $|H|$ 가 곡면의 각 점에 있어서 곡률을 표시하는 것으로 이용되는데, 곡면의 각점에 있어서 최대 곡률과 최소 곡률의 적으로 정의된 가우스, 곡률들을 계산한 이 방법은 영상의 노이즈의 영향을 줄이고 특징점의 위치가 어긋나는 것을 최소화 하기 위한 과정이다. 식 (2)는 가우스 곡률이다. 즉, 특징점 추출은 식 (2)를 이용한다.

$$k = \frac{|H|}{(1 + I_x^2 + I_y^2)} \quad \text{식 (2)}$$

그림 2. 는 실험에 사용된 입력영상과 특징 점이 추출된 영상이다

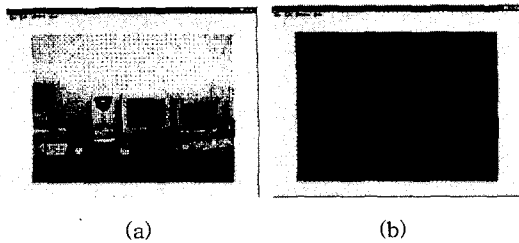


그림 2. 입력영상(a) 과 특징점 이 추출된 영상(b)

2.2 영상 분할(Image Segmantation)

영상 분할은 입력 영상을 동질성(homogeneity)과 연결성(connectivity)을 만족하는 영역들로 나누는 것으로 정의 할 수 있다. 여기서 동질성은 영역내의 모든 화소들이 공통의 성질을 갖는 것이고, 연결성은 영역내의 임의의 두 화소 사이에 연결된 경로(path)가 존재하는 것을 나타낸다.

본 논문에서는 명암도의 유사성에 착안한 동질성의 영역내의 화소에대한 영역을 우선 전처리로써 Median Filtering 으로 영상의 잡음을 제거하고, 영상의 윤곽선을 추출 후 레이블링을 통한 각 영역을 분리한다.

2.3 특징점 정합 (Feature Points Matching)

2.3.1 기존 연구

영상 안에서의 특징 점들에 대해서만 정합을 수행하는 방법을 특징점 정합(Feature Point Matching)이라 한다. 특징 점 정합은 컴퓨터 비전의 여러 분야에서 중요한 역할을 하는데 스테레오 정합, 트래킹(tracking), 움직임 추정, 영상교정, 3차원 복원, 셀프 칼리브레이션 등을 그 중요한 응용분야로 들 수 있다. Deriche[4]는 특징 점 사이의 정합 정도를 측정하는데 여러 가지 평가 기준을 사용하였다. 따라서, 정합정도를 수치적으로 나타내기가 어려우며 정합 가능한 특징 점을 찾아내는데 오히려 혼동을 증가시킨다. Zhang은 초기 정합의 결과로 발생하는 애매성의 해결을 위해 이완화(relaxation) 기법을 이용하였는데 이는 성능에 비해 많은 시간을 요구한다는 단점을 가지고 있다. 하지만, 위에서 제시한 단지 정합에 관한 연구는, 영상을 모자이크하는 데는 적합하지 않다. 우선, 많은 시간과 계산량을 요구하고, 영상의 회전과 크기에 대해서는 적용하는데 충분하지 않으며, 최적의 특징 점을 선택하여 모자이크하는 데는 불필요한 단계를 가지고 있다.

2.3.2 Hausdorff Distance 기반 정합(Matching)

Hausdorff distance는 두 영상을 비교하는데 주로 이

용되는 방법으로 분할된 첫번째 영상 영역에서 두번째 분할된 영상의 영역을 계층적으로 비교하여 Hausdorff distance가 최대인 점을 측정한다. 각각의 영역에 대해서 반복으로 최적의 4개의 대응점을 찾는다. 식 (3)은 Hausdorff Function 이고, 그림 3은 매칭을 시도한 과정을 보여주고 있다.

$$h(A, B) = \max_{a \in A} \left\{ \min_{b \in B} \{d(a, b)\} \right\} \quad \text{식 (3)}$$

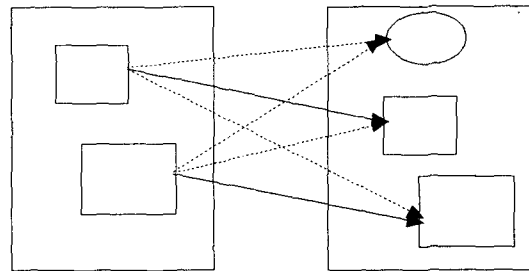


그림 3. Hausdorff Distance 를 이용한 매칭 절차

2.4 영상 변환과 합성

동일한 시점에서 촬영된 두 장의 영상 위치는 2차원 투영변환행렬 (M)에 의해서 결정된다. 변환행렬 M 은 다음과 같이 정의한다.

$$M = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{bmatrix} \quad \text{식 (4)}$$

식 (4)의 증명은 [1]에 있다.

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix} \quad \text{식 (5)}$$

여기에서, (x, y) 는 영상에서 최적의 특징 점으로 선택된 픽셀 좌표이고, (x', y') 는 대응되는 영상에서 픽셀 좌표이다. 위 식의 관계를 이용하여 식 (6)와 같은 방정식으로 앞에서 구한 4쌍의 초기 대응점을 구할 수 있다.

$$\begin{aligned} x' &= \frac{m_0x + m_1y + m_2}{m_6x + m_7y + 1}, \\ y' &= \frac{m_3x + m_4y + m_5}{m_6x + m_7y + 1}, \end{aligned} \quad \text{식 (6)}$$

본 논문에서는 비선형 이송 오차 최소화 기법으로 Levenberg-Marquardt 알고리즘을 사용하였다. 이 방법은 중첩영역 안에서 최소화되는 여러 e_i 를 구해 새로운 최적의 변환 행렬 M 을 결정하는 방법으로 알려지지 않은 운동 요소 $\{m_0, \dots, m_7\}$ 의 함수인 e_i 의

편미분 값을 이용한다. e_i 는 식 (7)과 같이 정의한다.

$$e_i = I'(x_i, y_i) - I(x_i, y_i) \quad \text{식 (7)}$$

최적의 변환 매트릭스로 영상 재구성시 각각 명암도가 다른 두 영상에서 재구성시 경계선이 발생하게 된다. 이를 제거하기 위해서 각 영상에 가중 평균 (weighted average) 알고리즘을 이용하여 눈에 보이는 경계선을 줄인다. 식(8)은 x 와 y 방향에 대한 blending function 이다.

$$I_b(x, y) = \frac{w_A(x, y)I_A(x, y) + w_B(x, y)I_B(x, y)}{w_A(x, y) + w_B(x, y)}$$

식 (8)

식 (9)는 weight function 이다.

$$w(x', y') = w_x(x')w_y(y') \quad \text{식 (9)}$$

결국, 식 (10)을 이용하여 최종 영상 모자이킹을 수행한다. 여기서 I' 는 명암도에 대한 응답 값을 나타낸다.

$$I' = \frac{\sum_k w_k(x_i, y_i)I_k(x_i, y_i)}{\sum_k w_k(x_i, y_i)} \quad \text{식 (10)}$$

3. 실험 및 고찰

실험은 제약이 없는 카메라로부터 획득한 영상으로, 실험에 사용된 영상은 모두 하나의 카메라로부터 약간의 시간 간격을 두고 획득한 영상이고, 전처리를 통해 256 그레이레벨을 갖도록 했다. 촬영장비는 Kodak DC120 디지털 카메라이고, 영상 모자이킹에 관한 실험은 Pentium 500MHz 와 128MB의 메모리를 장착한 PC 환경에서 진행하였다. 사용 OS는 Windows 98이다.

3.1 실험 결과

그림 4. 는 회전된 영상에 대해 모자이킹을 수행한 결과의 영상에 대한 결과를 보여주고 있다.

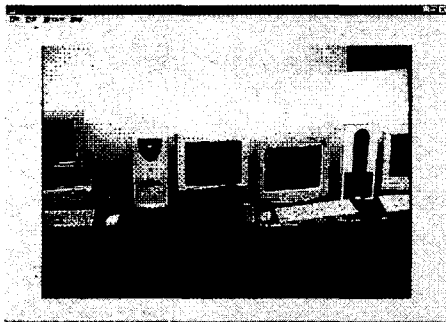


그림 4. 최종 모자이크 영상

3.2 고찰

본 실험에서에서의 계산 량이 많다는 것을 단점으로 들 수 있다. 영역안의 특징점을 매칭하기 위해서 소요 되는 시간과 최적의 변환 행렬을 구하기 위해서는 많은 시간이 필요하다. 영역을 보다 정확하게 분리한다면 계산 시간을 감소 할 수 있고, 제안하는 방법은 자유로운 카메라로부터 획득한 두장의 영상을 인간의 개입없이 특징 점을 찾아 특징 점에 대한 매칭 후 자동적으로 모자이크 영상을 생성하는데, 의의를 가질 수 있다.

4. 결론

본 논문에서는 자유로운 카메라로부터 약간의 시간차이를 두고 획득한 두장의 영상을 하나의 영상으로 인간의 개입없이 자동적으로 모자이킹 영상을 얻는 방법에 대해 기술하였다. 먼저, 영상의 특징 점과 영상의 영역을 분리하여 각각의 영역에 대한 특징점을 이용해 매칭을 시도하였고, 최적의 매칭된 4쌍을 얻어 초기 변환행렬을 비선형 이승 오차 최적화 알고리즘에 적용해 최적의 변환 행렬을 구한 후 두 영상을 합성하였다. 합성중에서 발생 할 수 있는 불연속 경계선은 가중평균 (Weighted average)를 적용하여 제거하였다.

참고문헌

- [1] R. Szeliski, "Video mosaics for virtual environments", IEEE Computer Graphics and Applications, P22-30, March 1996
- [2] R. Szeliski, "Image Mosaicing for Tele-Reality Applications", CRL Technical Report 94/2, Digital Equipment Corporation, 1994
- [3] I. Zoghlami and O. Faugeras and R. Deriche, "Using geometric corners to build a 2D mosaic from a set of images", INRIA, 2004 route des Lucioles, BP93
- [4] Thierry Blaszk, Rachid Deriche, "Recovering and Characterizing Image Features Using An Efficient Model Based Approach", INRIA, N2422, 1994
- [5] Naoki Chiba, Hiroshi Kano, Minoru Higashihara, Masashi Yasuda, and Osumi, "Frature-Based Image Mosaicing", MVA'98 IAPR Workshop on Machine Vision Applications, pp5-10, 1998
- [6] W. Puech, A. G. Bors, J.M. Chassery, I. Pitas, "Mosaicing of Paintings on Curved surfaces," IEEE Workshop on Applications of Computer Vision, Sarasota, USA, pp.50-55, 1996
- [7] W. H. Press, B. P. Flannery, A. A. Teukolsky, and W. T. Vetterling, "Numerical Recipes in C: The Art of Scientific Computing" Cambridge University Press Cambridge, England, Second edition, 1992