

트리를 이용한 이미지 모자이크 시스템 구축

이지현* · 양원보* · 전근환**

*군산대학교

**군장대학 경영 전산학부

Image Mosaic System Construction Using Tree

Ji-hyun Lee* · Won-bo Yang* · Keun-hwan Jeon**

*Kunsan University

**Faculty. of management & Computer Science, Kunjang College

E-mail : jklee@k2.kunsan.ac.kr

요약

이미지 모자이크 시스템이란 일련의 여러 영상, 혹은 동영상을 하나로 정합하는 것을 일컫는다. 본 논문에서는 일반적이고 직접적인 방법에 의한 이미지 모자이크 구축 방법을 피하고, 효율적으로 이용할 수 있도록 트리를 이용하여 모자이크 시스템을 구축하는 것을 제안한다. 또한, 정적 이미지 모자이크 및 동적 이미지 모자이크를 구축하는 방안에 대해서 제시한다.

Abstract

Image mosaic system is called that switching several image, or moving image into a image. In this paper, we represented that not image mosaic construction with respect to traditional and directional method but representing which mosaic system construction about using tree. Also, in this paper represent study that static image mosaic and dynamic image mosaic.

1. 서론

영상과 그래픽에 대한 수많은 발전에 힘입어 인간들은 더욱 더 영상에 대해 실세계를 보는 것과 같은 느낌을 원하게 되었으며, 그러한 요구로 인해 여러 영상들을 하나의 영상으로 생성하는 모자이크 영상이 발전되었다. 모자이크 영상은 연속된 비디오에 포함된 많은 정보들을 한 영상으로 여러 영상들의 내용을 수용하게 합성하여 표현한 것이다. 모자이크 영상 형성의 부류는 배경중심의 정적 모자이크, 동적 객체 표현을 위한 동적 모자이크, synopsis 모자이크로 분류된다. 모자이크 구축 과정의 기술들은 연속된 영상의 정렬, 통합, 중복성에 대한 분석으로 이루어진다. 비디오 압축, 확장된 시각화, 비디오 향상 등을 포함하는 비디오 색인화, 탐색, 조정에서 모자이크 표현은 강력한 비디오 응용기술이다. 파노라마 모자이크 시스템에 대한 응용분야로는 상당히 많이 존재하는데, 그 중에 가장 전통적인 분야는 공중 인공위성 사진의 구축이고, 최근의 연구 분야는 장면 고정과 장면 감지, 비디오 압축과 색인, 그리고 카메라의 해상도이며, 심지어 간단한 사진

편집까지도 연구되고 있다.

모자이크 영상의 응용분야는 상당히 광범위하게 이용되어 지고 있다. 그 중 가장 전통적인 응용 분야가 공중 인공위성 사진의 구축이고, 최근의 연구 분야는 장면 고정과, 비디오 압축과 인덱싱, 그리고 카메라의 해상도와 사진 편집등이 있으며, 현재 가장 활발히 진행되는 연구 분야는 가상 현실에 이용하는 모자이크 영상이다.

본 논문에서 제안하는 트리 기반 모자이크 영상은 모자이크 영상 생성을 위해 연속적인 영상들을 계속 적으로 비교해 점차적인 하나의 모자이크 영상을 만들어 가는 것이 아니라, 두 개의 영상을 비교해 하나의 부분

적인 모자이크를 생성하고 또한 그 다음 단계에서 이런 부분적인 모자이크 영상을 좀 더 큰 모자이크로 만들고 최종적인 상위 단계에서 총체적인 모자이크 영상을 생성하는 것이다. 이런 트리 기반 모자이크 영상은 직

접적인 모자이크 영상 생성 기법 보다 더욱 빠르게 생성 할 수 있으며, 또한 트리 기반 모자이크 영상을 가지고서 원 영상을 재현하는 방법에 대해서 제시한다.

II. 관련 연구

현재 상당히 많은 논문에서 optical flow 및 차 영상 그리고 영상의 좌표 변환 기법에 대해서 이야기하고 있다. 또한, 많은 논문들이 영상 정렬과 움직임 측정에 대한 문제들을 제시하였다. 이 장에서는 도자이크 영상 생성을 위한 기존의 방법들을 재조명해본다.

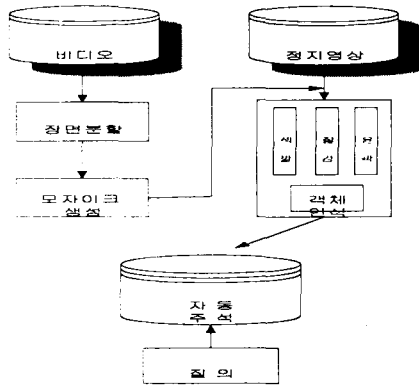


그림 1. IBM사의 ImageMiner 시스템의 개략도

모자이크 시스템의 활용으로는 IBM과 APPLE사에서 각각 시스템들을 개발하였고, 계속적으로 진행 중에 있다. 또한, 외국의 각 연구단체를 비롯하여 대학내에서 모자이크 시스템을 개발을 촉구하려는 방안에 대해 모색중이다. IBM사의 ImageMiner라는 시스템은 IBM사의 트레이드마크로서 비디오 분석과 정지영상을 합한 시스템이다. ImageMiner 시스템은 비디오로부터 입력받은 영상들을 가지고서 장면 분할을 하고 모자이크 기술을 이용하여 하나의 새로운 모자이크 영상을 생성한다. 이런 모자이크 영상을 분석하는 단계로서 color(색), texture(질감), contour(윤곽) 정보를 이용하며, 또한 객체 인식을 통해서 모자이크 영상의 내용들을 저장한다. 이렇게 저장된 내용물들을 가지고서 자동적인 주석을 달아 사용자의 질의에 응하는 시스템이다. 위의 그림1이 이를 개략적으로 표현해 준 IBM사의 ImageMiner 시스템의 개략적인 흐름도를 나타낸다. [14]

모자이크 영상을 생성함에 있어 가장 일반적인 가정은 영상들 사이에서의 좌표 변환은 단지 이동되었다는 점이다. 이렇게 단지, 이동되어진 영상들은 최소한의 제약으로 간단히 구현되어질 수 있고, 생성되어진 모자이크 영상은 질이 높은 고해상도의 모자이크가 생성될 수 있으며, 계산 시간 또한 빠른 이점이 있다. [2] 하지만, 이런 방법은 카메라의 신축, 회전, 팬(pan), 틸트(tilt) 같은 커다란 변화를 제어하기가 쉽지 않은 문제점을 가지게 된다.

일반적인 영상 좌표 변환 모델들은 표1과 같다.

표 1. 영상좌표 변환들

모델	X에서 X'로의 좌표 변환	파라미터
이동	$X'=X+b$	$b \in R^{2 \times 1}$
Affine	$X'=AX+b$	$A \in R^{2 \times 2}, b \in R^{2 \times 1}$
양선형	$x'=q_{x'xy}XY+q_{x'x}X+q_{x'y}Y+q_{x'}$ $y'=q_{y'xy}XY+q_{y'x}X+q_{y'y}Y+q_{y'}$	$q_i \in R$
원근 투영	$X' = \frac{AX+b}{C^T X+1}$	$A \in R^{2 \times 2}, b, C \in R^2$
Pseudo-perspective	$x'=q_{x'x}x+q_{x'y}y+q_{x'x^2}+q_{x'xy}$ $y'=q_{y'x}x+q_{y'y}y+q_{y'xy}+q_{y'y^2}$	$q_i \in R$

영상의 변환 중에 affine model(회전, 이동, 신축)의 부분 집합으로 고려하여 모자이크 영상을 생성하기도 한다[3]. 정적인 배경에 대해, 이런 affine model은 카메라의 줌(zoom), 카메라의 시각 축에 대한 회전, 이동의 변화에 대해서 정확히 계산될 수 있다. 하지만 affine model은 카메라의 모든 변화를 알 수 없으며, 잘못된 매개변수들이 측정될 수 있다.

영상의 변환 모델 중 8-파라미터 원근 투영 변환은 가능한 카메라의 모든 변화들을 정확하게 고려하여 바람직한 8개의 매개변수들을 계산할 수 있다. 하지만 이런 8-파라미터는 상당히 수학적이며 매개변수를 발견하기 위한 계산이 어렵다는 단점을 가지고 있다.

8-파라미터 양선형 모델은 영상 처리, 의학 영상, 원거리 측정, 컴퓨터 그래픽등 광범위한 분야에서 널리 이용되고 있으며, 8개의 정확한 파라미터를 계산할 수 있다. 이런 8개의 파라미터 모델에 대한 계산은 트레이드 오프(tradeoff)를 가지게 되는데, 잡음에 상당히 민감하게 반응을 한다.

III. 모자이크 기법

모자이크 영상을 생성하는 데 있어서, 우선적으로 선행해져야 되는 것이 모든 영상을 연속적으로 유사한 영상들끼리 정렬시켜야 한다는 것이다. 이는 비교할 두 영상간의 일치하는 지역을 하나의 부분으로 정합하는 것이 모자이크 영상이기 때문에 모든 영상들을 유사한 영상끼리 정렬하는 것이 먼저 선행되어야 한다. 유사한 영상끼리 연속적으로 정렬되었을 경우, 모자이크 영상을 생성하는데 있어서 더욱 빠르게 구축이 가능하다. 물론, 자동적인 정렬과 수동적인 정렬을 통해서 이루어 질 수 있다. 수동적인 정렬은 사람의 시각을 통해서 인지되어지는 감각을 통해서 이루어지고, 자동적인 정렬은 하나의 영상과 나머지 모든 영상을 비교함으로써 그 중의 가장 유사성이 큰 영상을 연속적으로 놓으면 된다. 그리고 나서 그 다음 영상을 기준으로 하고 또, 나머지 영상들을 비교함으로써 계산한다. 이런 방식을 계속적으로 행

함으로써 자동적인 정렬을 행할 수 있다.

3.1 직접적인 모자이크 영상 생성

입력되어진 모든 영상을 모자이크 영상으로 구축하는 방법 중에 가장 손쉬운 방법 중의 하나가 비 트리 기반 즉, 직접적인 모자이크 영상으로 구축하는 것이다. 처음 입력된 영상을 좀 더 커다란 영상(모자이크)안에 직접 넣어두고, 그 다음의 입력 영상과 일치하는 부분을 찾는다. 일치 지역을 찾게되면 그 영상을 모자이크 영상 안에 붙여 놓으면 된다. 연속적으로 입력되어지는 영상들을 계속해서 모자이크 영상내의 생성되어진 영상들과 비교해서 붙여 놓아 하나의 커다란 영상을 만드는 것이다. 이 방법은 간단하지만, 시간적인 소요가 많이 드는 단점을 가지고 있다. 직접적인 모자이크 영상 구축 방법의 그림은 아래와 같다.

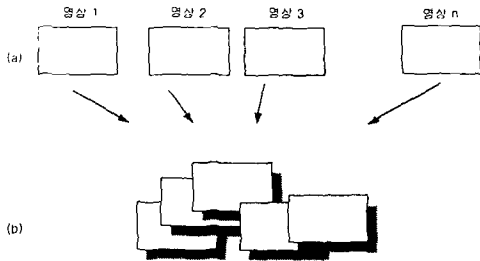


그림 2. 직접적인 모자이크 영상 생성

그림 2는 입력 영상 1을 모자이크 영상 (b)에 집어넣은 후 입력 영상 2와 비교를 해서 영상 1에다가 붙여 놓게되는 것이다.

3.2 트리 기반 모자이크 영상 생성

모자이크 영상 구축 기법 중의 또 다른 방법 중의 하나가 트리 기반 모자이크 영상 생성이다. 트리 기반 모자이크 영상 생성은 이웃하는 연속적인 영상들을 계속해서 비교함으로써 하나의 모자이크 영상을 생성하는 것이 아니라, 이웃한 단 두 개만을 비교하여 하나의 부분 모자이크 영상을 생성한다. 그리고, 그 이웃한 두 영상을 또 다른 부분 모자이크 영상으로 만들고, 모든 원 영상을 이런 부분 모자이크 영상으로 생성한다. 이렇게 두 영상을 하나의 부분 모자이크 영상으로 생성한 후, 부분 모자이크 영상들 중의 두 영상들을 비교하여, 다음 단계의 좀 더 큰 부분 모자이크 영상을 생성한다. 계속 반복함으로써, 총체적인 하나의 모자이크 영상으로 생성한다. 원 영상이 16개라면 3단계의 부분 모자이크가 존재하게 되며, 1단계에서는 8개의 부분 모자이크가 생성되고, 2단계에서는 4개의 부분 모자이크가, 3단계에서는 2개의 부분 모자이크가 생성되어진다. 그리고, 결국 그 두 영상을 합침으로써 총체적인 하나의 모자이크가 구축되어진다. 트리 기반 모자이크

영상 생성 방법의 그림은 아래와 같다.

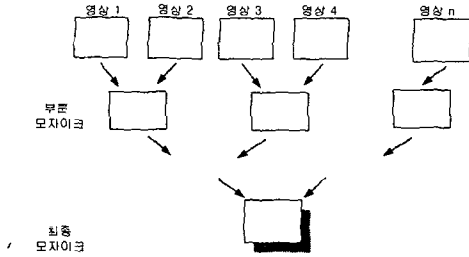


그림 3. 트리 기반 모자이크 영상

그림3은 입력되어진 영상1과 영상 2를 부분 모자이크 영상으로 생성하고, 영상 3과 영상 4를 또 하나의 부분 모자이크로 생성한다. 이렇게 모든 입력 영상을 가지고 부분 모자이크 영상으로 생성을 한 후에, 이런 부분 모자이크 영상들을 그 다음 단계의 부분 모자이크로 계속 적으로 생성하는 것이다. 트리 기반 모자이크 영상으로 생성하는 기법이 직접적인 모자이크 영상보다 계산 시간이 더욱 빠르게된다.

IV. 정적 모자이크 및 동적 모자이크 구축

모자이크 영상을 생성할 때에 동적 객체를 처리하지 않은 채 모자이크 영상을 구축하게 되면, 배경과 객체의 혼합된 영상으로 모자이크 영상이 모호하게 구축되어지게 된다. 그러므로 객체의 인식 및 추출은 모자이크 영상 구축을 위해 필수적인 요소가 되어진다. 이 장에서는 동적 객체를 추출하여 모자이크 영상에 이용하는 방안에 대해서 제시한다.

4.1 정적 모자이크

기본적인 배경 영상은 변하지 않고, 즉 카메라의 이동이 없이 움직이는 객체가 있는 경우에, 동적 객체를 추출하여 구축되어진 배경 모자이크 영상 위에다 표현함으로써 새로운 모자이크 영상을 생성한다. 배경 영상이 변하지 않고 객체의 움직임을 추출하기 위해서는 일반적으로 차 영상 기법을 이용한다. 즉, 두 영상 기준 영상과 비교 영상의 화소 대 화소간의 차 연산을 이용하여 계산한다. 차 영상의 공식은 식 1)과 같다.

$$D(x, y) = |I_a(x, y) - I_b(x, y)| \quad (1)$$

얻어진 값 D를 가지고 기준 임계치 보다 크면 객체의 이동이 있는 영상으로 간주한다. 여기서, 객체의 부분을 탐지하기 위해서는 지역적인 탐색이 필요하다. 본 논문에서는 지역적 탐색을 위해 16 × 16 매크로 블락(macro block)을 이용하였다. 식2)와 같다.

지역적 차 연산(Local Difference)은 기준영상 I_a 에서 참조 영상 I_b 를 뺀 값으로 이를 16 × 16 화

$$LD(x, y) = \sum_{y=1}^M \sum_{x=1}^N |I_a(x, y) - I_b(x, y)| \quad (2)$$

소의 크기만큼 계산한다. 객체의 추출을 간단하게 하기 위해 직사각형으로 분리하는 방법과 정확한 추출을 위해 객체의 위치를 이용하여 정확히 뽑아내는 방법이 있다.

4.2 동적 모자이크

카메라가 이동하는 경우, 즉 배경이 움직이게 될 때 모자이크 영상을 생성하면 배경 영상들이 합쳐져 커다란 새로운 영상이 구축되어진다. 일단 동적 객체 존재 시, 동적 객체를 추출하여 떼어놓은 후, 배경 모자이크를 구축한 후에 배경 모자이크 영상 위에 추출된 동적 객체를 표현해 준다.

첫째, 일단 배경 모자이크 영상을 구축하기 위해서는 두 영상간의 유사성을 찾아야 한다. 일반적으로 많이 이용되는 간단한 방법인 최소 사각형 오류 (Least sum of squared error) 기법은 식(3)과 같다.

$$E(C) = \sum_{b \in W} [I_i(X+b) - I_j(X+b+d_k)]^2 \quad (3)$$

둘째, 동적 모자이크 영상 생성을 위해서는 카메라의 움직임을 계산한 후에, 객체의 움직임을 인지해야 한다. 객체의 움직임을 위해서는 지역적 탐색이 이루어져야 한다. 지역적 탐색은 계산되어진 카메라의 움직임을 가지고서 두 영상간의 비교를 통해서 이루어지게 된다. 계산식은 식(4)와 같다.[13]

$$E(O) = \sum_{b \in W} [I_i(X+b) - I_j(X-u(k)+b+d_k)]^2 \quad (4)$$

여기에서 $u(k)$ 는 x축, y축으로의 방향 벡터이다. 비교되어지는 영상에서 방향벡터, 즉 카메라의 움직임의 거리만큼을 빼서 두 영상간의 에러값을 계산한다. 이 때 두 영상사이에서 객체의 움직임을 인지하는 임계치를 크게 주면 커다란 객체의 움직임만 인식되고, 임계치를 작게 주면 작은 움직임도 추출될 수 있다.

셋째, 배경 영상을 합성하는 것은 동적 객체를 추출한 후에 나머지 배경 영상만을 합성한다. 이 때, 추출된 동적 객체의 부분에는 나머지 다른 영상의 배경 부분이 삽입된다. 그림으로써, 동적 객체는 제거된 배경 모자이크 영상만을 생성할 수 있다.

마지막으로, 생성되어진 배경 모자이크 영상 위에 추출된 동적 객체를 원 영상의 위치에 맞게 표현해 줌으로써 이동하는 배경에 동적 객체를 시놉시스(synopsis)하게 나타내 주는 모자이크 영상을 구축할 수 있다.

5. 실험

본 논문에서 제안한 모자이크 시스템을 구현하기 위해서 SONY사의 HITACHI Digital Zoom 카메라로부터 영상을 입력받았다. 실험에 사용되

어진 영상은 320*240의 크기의 color image를 입력 영상으로 사용하였고, Pentium-450에서 Visual C++5.0을 이용하였다. 또한 정지 영상의 입력을 위해 여러 장의 사진도 사용하였다.



그림 4. 정적 입력 영상



그림 5. 정적 모자이크 영상

그림 4는 배경은 정지한 영상에 객체만 이동한 영상을 입력받은 것이고 그림 5는 이 세 영상을 정적 모자이크 영상으로 구축하여 객체의 이동 모습을 나타내 준 영상이다.



그림 6. 입력 영상



그림 7. 배경 모자이크 영상



그림 8. 동적 모자이크 영상

그림 6은 이동하는 배경을 통해서 동적 객체가 삽입되어 있는 영상이다. 입력 영상들을 가지고 배경 모자이크를 구축한 영상이 그림 7와 같다. 배경 모자이크 영상에는 각 입력 영상사이의 객체가 삭제되어 구축되었고, 동적 모자이크 영상은 배경 모자이크 영상에 동적 객체를 각 입력 영상의 위치에 알맞게 표현함으로써 표현한 것으로 그림 8과 같다.

6. 결론

본 논문에서는 모자이크 영상을 생성하기 위

한 기초적인 기법과 트리 기반 모자이크 영상과 정적 및 동적 모자이크 구축 방안에 대하여 제시하였다. 비트리 기반 모자이크, 즉 직접적인 모자이크 영상보다 계산 속도가 더욱 빠른 트리 기반 영상 생성을 통해서 모자이크를 구축하였다.

향후, 움직이는 객체에 대한 정확한 추출 및 객체의 연속적인 정보의 저장을 통해 모자이크 영상의 저장과 전송에 이용할 수 있는 방안에 대해 연구할 과제가 남아 있고, 카메라의 움직임이 복잡한 영상들에 대하여 더욱 정확한 기술들을 발전시켜야 할 것이다.

참고문헌

[1] Heung-Yeung Shum and Richard Szeliski, "Panoramic Image Mosaic", Microsoft Research, MSR-TR-97-23

[2] A. Tekalp, M.Ozkan, and M.Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," in Proc. of the Int. Conf. on Acoust., Speech and Sigt. Proc., (San Francisco, CA),pp. III-169, IEEE, Mar. 23-26, 1992.

[3] Q.Zheng and R.Chellappa, "A Computational Vision Approach to Image Registration," IEEE Transactions Image processing, July 1993. pages 311-325

[4] M.Irani and S.Peleg, "Improving Resoltuion by Image Registration," CVGIP, vol. 53, pp. 231-239, may 1991.

[5] L. Teodosio and W.Bender, "Salient video stills : Content and context preserved," Proc. ACM Multimedia Conf., 1993.

[6] G.Wolberg, Digital Image Warping. 10662 Los Vaqueros Circle, Los Alamitos, CA : IEEE Computer Society Press, 1990. IEEE Computer Society Press Monograph.

[7] Gulrukh Ahanger and Thomas D.C.Little, "A Survey of Technologies for Parsing and Indexing Digital Video", Journal of visual communication and iamge representation, Vol. 7, No.1, Maarch, pp.28-43, 1996, ARTICLE NO.00004

[8] Harpreet S.Sawhney and Serge Ayer, "Compact Representations of Videos Through Dominant and Multiple Motin Estimation", IEEEE, April 11, 1996

[9] Richard Szelisk, "Image Mosaicing for Tele-Reality Applications", Digital Equipment Corporation, Cambridge Research Lab, May, 1994, CRL 94/2.

[10] Richard Szeliski and James Coughlan, "Spline-Based Image Registration", Digital Equipment Corporation, Cambridge Research Lab, May, 1994, CRL 94/1.

[11] Richard Szelisk, Sing Bing Kang, "Direct Methods for Visual Scene Representation", June 24, 1995, IEEE Workshop.

[12] 양원보, 임문순, 이양원, "다중 영상에 의한 비디오 모자이크 시스템", 한국해양정보 통신학회 춘계종합학술대회, June, 1999, Vol.3 No.1, pp 104-108.

[13] 양원보, 김재원, 전근환, 이양원, "트리 기반 모자이크 영상과 재분배", 한국 정보 처리학회 추계학술발표대회, October 8, 1999

[14] J. Kreyβ, M.Roper, P.Alshuth, Th.Hermes, and O.Herzog, "Video Retrieval by Still Image Analysis with ImageMiner", SPIE Vol. 3022