

# 트리 기반 이미지 모자이크 시스템

김미호\*, 신성윤\*\*, 전근환\*\*, 이양원\*

\*군산대학교 컴퓨터정보과학과 · \*\*군장대학 컴퓨터응용학부  
e-mail:syshin@cs.kunsan.ac.kr

## Tree-based Image Mosaic System

Mi-Ho Kim\*, Seong-Yoon Shin\*\*, Keun-Hwan Jeon\*\*, Yang-Weon Rhee\*

\*Dept of Computer Science, Kunsan National University

\*\*Division of Computer Application, Kunjang College

### 요약

영상 내의 객체를 전체적으로 표현할 수 없거나 부분적으로 보이지 않는 부분을 갖는 일련의 여러 영상, 혹은 동영상상을 하나로 정합하여 확장된 영상을 나타내는 시스템을 이미지 모자이크 시스템이라 한다. 본 논문에서는 기존의 방법보다 효율적인 트리를 기반으로 한 모자이크 시스템을 구축하는 방법을 제안한다. 또한, 정적 이미지 모자이크뿐만 아니라 및 동적 이미지 모자이크를 구축하는 방안에 대해서도 제시한다.

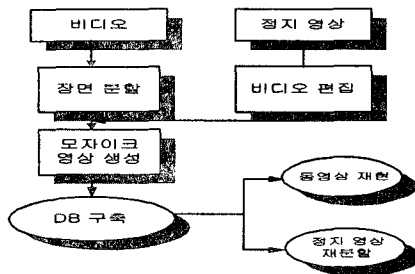
### 1. 서론

그래픽과 영상의 발전은 급속도로 빠르고 폭넓게 진행되어 하드웨어와 소프트웨어의 지원으로 실 영상과 더욱 유사한 영상을 볼 수 있게 되었고, 또한 사용자요구도 더욱 증대되었다. 모자이크 영상이라는 것은 연관되어지는 여러 비디오 및 정지 영상들을 합성하여 하나의 새로운 영상을 생성하는 것[1]을 말하며, 배경중심의 정적 모자이크, 동적 객체 표현을 위한 동적 모자이크, 시놉시스(synopsis) 모자이크로 분류된다[2].

모자이크 구축 과정의 기술들은 연속된 영상의 정렬, 통합, 중복성에 대한 분석으로 이루어진다. 파노라마 모자이크 시스템에 대한 응용분야로 가장 전통적인 분야인 공중 인공위성 사진의 구축과, 최근의 연구분야인 장면 고정과 장면 감지, 비디오 압축과 색인, 그리고 카메라의 해상도이며, 심지어 간단한 사진 편집까지 매우 다양한 분야에서 연구되고 있다 [2,3,4].

(그림 1)의 시스템을 보면, 입력 단계에 있어서 비디오 카메라로 찍은 동영상과 일반 카메라로 찍은 정지 영상을 갖는데 입력 영상이 일련의 비디오 영상이라면 장면 분할의 작업이 선행되어야 한다. 생성되어진 모자이크 영상을 원 영상으로 재현하게 되면 저장 및 전송측면에 있어 상당한 효과를 볼 수

있다[1,5]. 본 논문에서는 모자이크 영상을 생성하는 방안에 중점을 두고 있다.



(그림 1) 모자이크 시스템 개략도

본 논문에서는 입력영상을 트리 기반 모자이크 영상으로 생성하는 것과 동적 객체를 추출하여 배경 모자이크 영상으로 표현해주는 방법을 제시한다.

본 논문의 구성은, 2장에서는 관련 연구, 3장에서는 모자이크 영상 생성 시스템 및 트리 기반 모자이크 영상 기법, 4장에서는 정적 및 동적 모자이크 영상 구축 방안, 그리고 5장에서는 실험을 수행하고 6장에서 결론을 맺도록 한다.

### 2. 관련 연구

현재 상당히 많은 논문에서 optical flow 및 차 영

상, 영상의 좌표 변환 기법, 영상 정렬과 움직임 측정에 대한 연구를 수행하고 있는데. 모자이크 영상 생성을 위한 기존의 방법들을 살펴본다.

Shum & Szeliski는 입력 영상을 변환 행렬과 연관하여 모자이크 영상을 표현하는 방법[3]을 제안하였는데, 잡음에 매우 민감하다는 단점을 가지고 있다. Xiong과 Turkowski는 카메라를 90도로 회전하여 4개의 영상을 얻어 가상세계를 구축하는 방법[4]을 제시하였고, Michael Irani는 동적 모자이크[2]에 대하여 제안하였는데 동적 객체에 대해서 정확하게 외형을 추출하는 방법은 기술하지 않았다.

일반적인 영상 좌표 변환 모델들은 <표 1>과 같은데, 그 중 어파인 모델(affine mode)은 부분 집합으로 고려하여 모자이크 영상을 생성하는 방법이다 [6]. 정적인 배경에 대해 이런 어파인 모델은 카메라의 줌(zoom), 카메라의 시각 축에 대한 회전, 이동의 변화에 대해서 정확히 계산될 수 있다.

<표 1> 영상 좌표 변환들

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

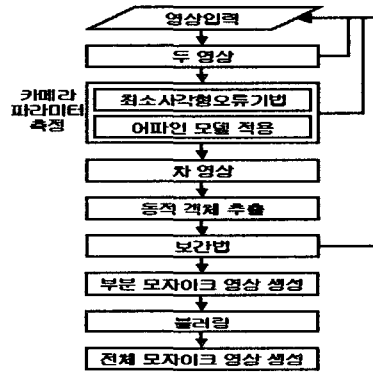
정확한 파노라마 구성을 위해서는 연속된 입력 영상이 기준 좌표계에 대해서 정확하게 정렬되어야 한다. 입력영상 자체에 배경이 움직이지 않을 경우에는 문제가 되지 않지만 카메라 동작이 포함된 입력 영상에 대해서는 카메라 동작 파라미터를 정확하게 추출한 뒤 보정 작업이 필요하다. 모자이크 영상을 생성에서 가장 일반적인 가정은 영상들 사이의 좌표 변환은 단지 이동되었다는 것이다. 단지 이동되어진 영상들은 최소한의 제약으로 간단히 구현될 수 있고 질 높은 고해상도의 모자이크가 생성될 수 있으며 계산시간이 빠르다는 이점이 있다[7]. 하지만, 이런 방법은 카메라의 신축, 회전, 팬, 틸트 같은 큰 변화를 제어하기 어렵다는 문제점을 갖는다.

### 3. 트리 기반 모자이크 영상 구축

모자이크 영상 생성에서, 먼저 선행될 사항은 모든 영상을 연속적으로 유사한 영상들끼리 정렬시키는 것으로, 이는 비교할 두 영상간의 일치하는 지역을 하나의 부분으로 정합하는데 필요하다.

#### 3.1 모자이크 시스템 흐름도

비디오 카메라로 찍은 영상과 일반 카메라로 찍은 영상이 입력으로 사용되는데, 일반 카메라로 찍혀진 영상은 각 영상들의 차이가 상당히 클 수 있다는 점이다. 이런 경우에 유사성 측정을 통하여 모자이크를 구축하는 것이 효과적이다.

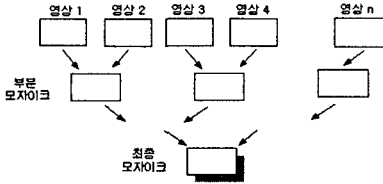


(그림 2) 모자이크 시스템

(그림 2)는 본 논문의 모자이크 영상 생성을 위한 시스템 흐름도로서, 카메라 영상들에서 두 영상씩 입력받아 사용한다. 두 영상간 카메라 이동 측정은 최소사각형오류기법 및 어파인 모델을 이용하며, 차 영상을 계산하고 동적객체를 추출하여 부분 모자이크 영상을 생성한다. 마지막으로 블러링을 적용하여 전체적으로 완전한 모자이크 영상을 구축한다.

#### 3.2 트리 기반 모자이크 영상 생성

트리 기반 모자이크 영상 생성은 이웃하는 영상들을 연속적으로 모자이크 영상을 생성하는 것이 아니라, 이웃한 단 두 개씩만을 비교하여 하나의 부분 모자이크 영상을 생성하는 것으로, 모든 원 영상을 이런 부분 모자이크 영상으로 생성한다. 이렇게 생성된 부분 모자이크 영상으로 새로운 부분 모자이크 영상 생성을 반복하여 총체적인 하나의 모자이크 영상으로 생성한다. 원 영상이 16개라면 4단계의 부분 모자이크가 존재한다. 트리 기반 모자이크 영상 생성 방법은 (그림 3)과 같으며 계산시간이 빠르다.



(그림 3) 트리 기반 모자이크 영상

#### 4. 정적 및 동적 모자이크 영상 구축

4.1 정적 배경과 동적 객체를 이용한 모자이크 구축  
 정적 배경에서 동적객체의 추출에는 차 영상 기법을 이용하는데, 두 영상의 화소 대 화소간의 차 연산을 이용하여 계산하는 것으로 식 (1)과 같다.

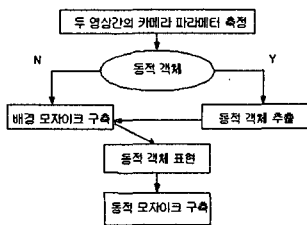
$$D(x,y) = |I_a(x,y) - I_b(x,y)| \quad (1)$$

여기서, 객체 부분을 탐지하기 위해서는 지역적인 탐색이 필요한데, 본 논문에서는 16 × 16 매크로블럭(macro block)을 이용하여 시 (2)와 같다

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

지역적 차 연산은 기준영상 I<sub>a</sub>에서 참조영상 I<sub>b</sub>를 뺀 값으로 이를 16×16 화소의 크기만큼 계산한다. 객체 추출은 직사각형으로 분리하는 방법 객체의 위치를 이용하는 방법이 있다.

#### 4.2 동적 배경과 동적 객체를 이용한 모자이크 구축



(그림 4) 동적 모자이크 구축

(그림 4)는 동적 모자이크 구축을 나타내는데 다음 4가지의 단계로 구성된다.

첫째, 두 영상간 유사성을 찾기 위해 정확한 카메라 파라미터를 추출해야 하는데, 가장 많이 이용되는 방법은 최소사각형오류기법으로 식 (3)과 같다.

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

식 (3)을 통해 가장 유사한 화소 값을 찾아내 이를 바탕으로 어파인 모델을 사용하여 영상의 회전, 신축 및 이동을 포함하는 카메라 파라미터를 측정한다.

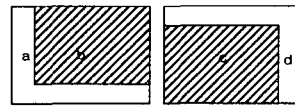
다. 어파인 모델은 식 (4)와 같다.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \quad (4)$$

둘째, 객체의 움직임을 인지해야 하는데 이를 위해서는 지역적 탐색이 이루어져야 한다. 이는 계산되어진 카메라 움직임을 가지고 두 영상간의 비교를 통해 이루어지며 방법은 식 (5)와 같다[1].

$$E(O) = \sum_{b \in W} [I_i(X+b) - I_j(X-u(k)+b+d_k)]^2 \quad (5)$$

여기에서 u(k)는 x축, y축으로의 방향 벡터로서 카메라의 움직임의 거리만큼을 나타내며 이를 때서 두 영상간의 에러 값을 계산한다. 이때 객체의 움직임을 인지하는 임계치를 크게 주면 커다란 움직임만 인식되고, 작게 주면 작은 움직임도 추출할 수 있다.



(그림 5) 동적객체판별을 위한 영역분할

(그림 5)의 영상은 b와 c의 부분은 두 영상간의 유사 지역을 표시한다라고 가정하면 동적객체를 추출하기 위해서는 다음과 같은 가정이 따른다.

첫째, a와 c, d는 일치하는 부분이 없다. 둘째, b는 c와 일치하고 d와는 일치하는 부분이 없다.

위와 같은 가정으로 a의 어느 부분과 c나 d의 어느 부분이 일치하면 이 지역은 객체가 된다. 또한 b와 d의 부분이 같게 되면 역시 동적객체임을 알 수 있다. 반대로, 두 영상간에 일치하는 b와 c간에 일치하지 않는 부분은 객체임을 판별할 수 있다.

셋째, 배경영상 합성은 동적객체추출 후 나머지 배경영상만을 합성한다. 이때, 추출된 동적객체의 부분에는 나머지 다른 영상의 배경 부분이 삽입된다.

마지막으로, 생성된 배경 모자이크 영상에 추출된 동적객체를 위치에 맞게 표현하여 동적객체를 시삽 시스템하게 나타내는 모자이크 영상을 구축한다.

### 제 5 장 실험

본 논문에서는 HITACHI Digital Zoom 카메라로 컬러 영상을 입력, 320\*240크기로 정규화 한 동영상과 여러 장의 정지 영상을 실험에 이용하였다.





(그림 6) 카메라로부터 얻어진 영상들

(그림 6)은 임의의 4개의 압력 영상이며, 이들로 구축된 파노라믹 모자이크는 (그림 7)과 같다.



(그림 7). 정적 모자이크 영상

(그림 7)은 (그림 6)의 정적 모자이크 영상이다.



(그림 8) 입력 영상



(그림 9) 정지한 배경에서 동적 모자이크 영상

(그림 8)은 배경이 정지 영상에 객체 이동 영상을 입력받은 것이고, (그림 9)는 이들을 모자이크 영상으로 구축하여 객체 이동을 나타내 준 영상이다.



(그림 10) 입력 영상



(그림 11) 동적객체를 제거한 배경 모자이크 영상



(그림 12) 동적 모자이크 영상

(그림 10)은 동적배경을 통해서 동적객체가 삽입

되어 있는 영상이고 (그림 11)은 입력 영상들을 가지고 객체가 삭제된 배경 모자이크를 구축한 영상이다. 동적 모자이크 영상은 배경 모자이크 영상에 동적객체를 표현한 것으로 (그림 12)와 같다.

### 제 6 장 결론

본 논문에서는 트리 기반 모자이크 영상을 생성하는 과정에서 정적 및 동적 모자이크 영상을 구축하는 방안을 제시하였다. 정적 모자이크 영상에 대해 차 영상과 지역적 차 연산 기법을 통해 정적 모자이크 영상을 생성하였으며, 동적 모자이크 영상을 구축하기 위해 동적객체를 제거하여 배경 모자이크 영상을 생성하고 그 위에 동적객체를 표현하여 시남시한 동적 모자이크 영상을 생성하였다.

카메라의 이동이 큰 경우의 모자이크 영상 생성과 웹 상에서 구현 방안에 대하여 지속적인 연구가 필요하다고 사료된다.

### 참고문헌

- [1] Wonbo Yang, Jaewon Kim, Keunhwan Jeon, Yangweon Rhee, "Tree-based mosaic image and redistribution", Proc. KIPS, Oct. 8, 1999
- [2] M.Irani, P.Anandan, S.Hsu, "Mosaic based representation of video sequence and their applications", IEEE Proc. ICCV'95, 605-611.
- [3] Heung-Yeung Shum and Richard Szeliski, "Panoramic Image Mosaic", MSR-TR-97-23
- [4] Y.Xiong, K.Turkowski, "Creating image-based VR using a selfcalibrating fisheye lens", Proc. CVPR' 97, 237-243.
- [5] Wonbo Yang, Moonsun Lim, Yangweon Rhee, "Video Mosaic System by Multi-Image", Proc. KIMICS, Jun, 1999, Vol.3 No.1, pp 104-108.
- [6] Q.Zheng and R.Chellappa, "A Computational Vision Approach to Image Registration," IEEE Trans. Image Processing, Jul 1993. pp 311-325
- [7] A. Tekalp, M.Ozkan, and M.Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," Proc. Int. Conf. on Acoust., Speech and Sigt. Proc., (San Francisco, CA),pp.III-169, IEEE, Mar. 23-26, 1992.