

# 카메라와 동적 객체의 파라미터 측정을 이용한 정적 및 동적 모자이크 구축

양원보\*, 임문순\*\*, 최효성\*\*\*, 이양원\*

\*군산대학교 컴퓨터 정보 과학과

\*\*충남 시초 초등학교

\*\*\*전북 군산 진포 중학교

e-mail: bbong@cs.kunsan.ac.kr

## Static and Dynamic Mosaic Construction using Parameter of Camera and Dynamic Object

Won-bo Yang\*, Moonsun Lim\*\*, Yang-weon Rhee\*

\*Dept of Computer Information Science, Kunsan University

\*\*Sicho Elementary school, Choongnam

\*\*\*Kunsan Jinpo Middle school, Chonbuk

### 요약

비디오는 상당히 많은 정보를 함축적으로 포괄하는 있는 자원이다. 하지만 비디오의 내용물을 처리하고 이해하기에는 많은 어려움이 있다. 모자이크라고 하는 것은 연속적인 영상들로부터 정합 하여 하나의 새로운 영상을 생성함을 일컫는다. 그러므로 모자이크 영상을 통해서 비디오의 많은 정보들을 효과적으로 표현해 줄 수 있다.

본 논문에서는 모자이크 구축을 위해 이동하는 카메라를 통해 입력을 받아 모자이크 영상을 생성하는 방안을 제시한다. 정적 모자이크 구축을 위해 두 영상간의 방향 벡터만큼 이동 후 차 영상을 계산하여 구축하는 방안을 제시한다. 동적 모자이크 구축을 위해 움직이는 카메라와 동적 객체의 파라미터를 모호성을 방지하기 위한 알고리즘을 제시하고, 동적 객체 관별을 위한 영역 분할 기법에 대하여 제시한다.

### 1. 서론

최근 멀티미디어의 급속한 발전으로 인해 멀티미디어의 시대가 도래하게 되었고, 많은 사용자들이 이용하고 있는 실정이다. 하지만 멀티미디어의 막대한 자료의 처리 및 전송 그리고 정보의 이해가 실로 어려운 문제로 남아있다. 특히, 비디오는 많은 정보를 함축적으로 포괄하고 있는 것으로서, 사용자가 그 비디오의 내용물을 이해하기 위해서는 처음부터 끝까지 탐색해 나가야만 되는 부담을 주게된다.

관련 있는 여러 영상들을 하나의 영상으로 새롭게 생성하는 것을 '모자이크'라고 부른다. 모자이크 영상에는 비디오의 한 장면

에 있어서 많은 정보들을 놓칠 수 있는 문제점을 제거해준다. 모자이크 영상은 정지 영상과 비디오로부터 입력을 받게된다. 입력되어지는 영상이 정지 영상이라면 자동 정렬에 의해 관련 있는 영상들끼리의 정합을 유도하고 비디오 영상이 입력되어지면 이웃하는 영상을 정합하기 위해 먼저 선행되어야 할 장면 분할이 필수적으로 요구되어진다. 하나의 장면을 대표하는 프레임을 대표 프레임이라 하는데, 이 대표 프레임만을 가지고서 장면을 모두 이해하기에는 어려움이 많이 존재한다. 하지만, 모자이크 영상은 이웃하는 영상을 모두 포괄하고 있게 되므로, 하나의 장면을 총체적으로 이해시킬 수 있다. 모자이

크 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. 본 논문에서는 이동하는 카메라를 가지고 입력을 받아 움직이는 카메라의 파라미터 측정 및 동적 객체의 파라미터를 측정하여 모자이크 영상을 구축하는 영역 분할 방법에 대하여 제시한다.

## 2. 관련연구

모자이크 시스템을 구축하기 위한 논문은 커다란 이슈를 불러일으키며 많은 연구가 진행되어 왔으며, 더욱 발전하고 있는 실정이다. 이 장에서는 모자이크 시스템을 구축하기 위한 기존의 방법들을 재조명해 본다. 현재 상당히 많은 논문에서 optical flow 및 영상의 좌표 변환 기법을 이용하고 있다. 또한, 많은 논문들이 영상 정렬과 움직임 측정에 대한 문제들을 제시하였다.

일반적인 평면에 투영하는 방법이 아닌 각 입력 영상을 변환 행렬과 관련하여 모자이크 영상을 표현하는 것에 대하여 Shum & Szeliski는 제안하였다. 하지만 이 방법은 잡음에 매우 민감하다는 단점을 가지고 있다 [1]. 카메라를 90도로 회전하여 4개의 영상을 얻어 가상세계를 구축하는 방안을 Xiong과 Turkowski는 제시하였다[5]. 또한 동적 모자이크에 대해서 Michael Irani는 제안을 하였지만, 하지만 움직이는 객체에 대해서 정확하게 외형을 추출하는 방법은 기술하지 않았다.[4]

그림 1은 본 논문의 전체적인 모자이크 시스템 구조를 나타내고 있다. 입력되어지는 영상이 비디오라면 먼저 장면 분할의 작업을 거친다. 모자이크라고 하는 것이 이웃하는 영상들 속에서 유사성을 측정하여 하나의 영상을 만드는 것이기 때문에 필수적으로 선행되어야 할 장면 분할의 과정을 거쳐야만 한다. 또한 입력 영상이 정지 영상이라면 가

장 유사율이 높은 영상을 이웃하게 하여 자동적으로 정렬하는 작업을 거친다.

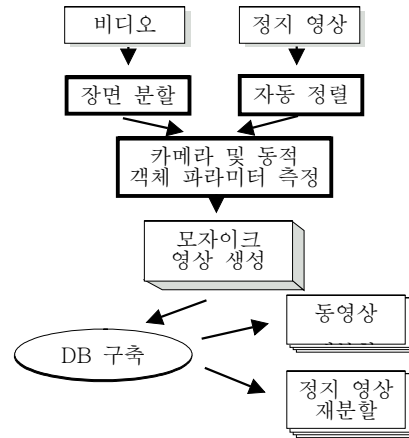


그림 1. 전체적인 모자이크 시스템 구조

이렇게 각 입력 자료에 따라 선행되어야 할 장면분할과 자동 정렬의 작업이 끝나게 되면, 움직이는 카메라 및 동적 객체의 파라미터 측정을 작업을 거친 후 모자이크 영상을 생성한다. 본 논문에서는 전체적인 시스템 중에서 카메라 및 동적 객체의 파라미터를 측정하는 부분과 모자이크 영상을 생성하는 과정에 대하여 자세히 기술할 것이다.

## 3. 정적 모자이크 구축

### 3.1 카메라 파라미터 측정

모자이크 영상을 구축하기 위해 선행해야 하는 것이 입력 영상 물이 비디오 영상이라면 하나의 샷으로 분할해야하며, 정지 영상이라면 모든 영상을 연속적으로 유사한 영상들끼리 정렬시켜야 한다는 것이다. 이는 비교할 두 영상간의 일치하는 지역을 하나의 부분으로 정합 하는 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 인지

되어지는 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 정렬한다. 이런 방식을 계속적으로 행함으로써 자동적인 정렬을 행할 수 있다.

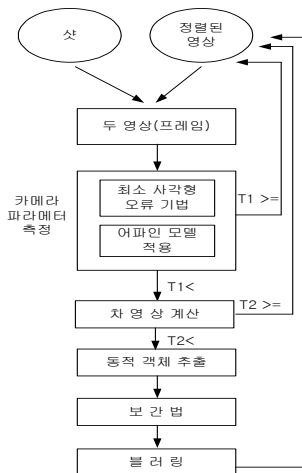


그림 3 모자이크 영상 구축을 위한 흐름도

그림 2는 모자이크 영상 구축을 위한 흐름도를 간략히 보여준다. 간략하게 살펴보면, 카메라로부터 찍혀진 실세계의 영상들을 분할되어진 샷으로써 입력받고, 정지 영상이라면 자동 정렬된 영상을 받아들여 두 영상(프레임)을 비교한다. 두 영상간의 카메라 이동을 측정하기 위해 최소 사각형 오류 기법 및 어파인 모델을 이용한다. 두 영상의 차 영상을 계산하고 동적 객체의 존재 시 동적 객체를 인지하여 추출한 후, 보간을 시켜서 모자이크 영상을 구축한다. 마지막단계는 블러링 기법을 적용하여 전체적으로 완만한 모자이크 영상을 구축한다.

### 3.2 어파인 모델

일단 두 영상간의 유사성을 찾기 위해서 정확한 카메라 파라미터를 추출하여야 한다. 카메라 파라미터를 위해 일반적으로 가장 많이 이용되는 간단한 방법은 최소 사각형 오

류 (Least sum of squared error) 기법으로 식 (1)과 같다.

$$E(C) = \sum_{b \in W} [I_i(X+b) - I_j(X+b+d_k)]^2 \quad (1)$$

식 (1)에서 X는 x,y의 픽셀 위치를 말하고, b는 한 영상의 정사각형의 윈도우를 일컫는다. 기준영상 I<sub>i</sub>에서 I<sub>j</sub>의 영상의 차이 값의 제곱으로 계산되어지는데 참조영상 I<sub>j</sub>의 모든 d<sub>k</sub>의 블록을 통하여 최소 사각형 에러 값을 계산한다. 이 때 계산되어진 사각형 에러 값 중 최소의 값을 선택한다.

식 (1)을 통해서 가장 유사한 화소 값을 찾아낸 후 이를 바탕으로 어파인 모델을 사용하여 영상의 회전, 신축 및 이동을 포함하는 카메라 파라미터를 측정한다. 어파인 모델을 자세히 기술하면 식(2)와 같다.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \quad (2)$$

어파인 모델을 계산하기 위해서는 기준 영상과 참조 영상의 비선형의 세 개의 점들 (x<sub>1</sub>, y<sub>1</sub>, x<sub>2</sub>, y<sub>2</sub>, x<sub>3</sub>, y<sub>3</sub>), (x<sub>1</sub>, y<sub>1</sub>, x<sub>2</sub>, y<sub>2</sub>, x<sub>3</sub>, y<sub>3</sub>)이 필요하다.

$$x' = ax + by + e \quad y' = cx + dy + f \quad (3)$$

이 세 개의 점들을 통하여 세 개의 선형 방정식을 만들어 어파인 모델의 6개의 파라미터(a,b,c,d,e,f)를 계산한다. 기준 영상의 세 점과 참조 영상의 세 점을 통해서 세 개의 선형 방정식을 계산하여 6개의 파라미터를 계산한 후 참조 영상의 모든 픽셀에 대응하는 점들을 배치한다.

### 3.3 정적 모자이크 영상 생성

정적 모자이크 영상이라고 하는 것은 카메라의 움직임은 있으나, 동적 객체가 없는 경우를 말한다. 즉, 움직이는 객체가 존재하지 않는 배경만을 찍은 영상을 일컫는데, 동적 객체의 존재 여부는 쉬운 차 영상을 통하여 계산한다. 차 영상의 공식은 식 (1)과 같

다.

$$D(x, y) = |I_a(x, y) - I_b(x, y)| \quad (4)$$

얻어진 값  $D$ 를 가지고 기준 임계치 보다 크면 객체의 이동이 있는 영상으로 간주한다. 여기서, 객체의 부분을 탐지하기 위해서는 지역적인 탐색이 필요하다. 본 논문에서는 지역적 탐색을 위해  $16 \times 16$  매크로 블록(macro block)을 이용하였다. 모든 매크로 블록이 임계치 이하이면 두 영상간의 동적 객체가 존재하지 않는 것으로 간주하여 정적 모자이크 영상을 구축한다. 임계치 이상인 블록이 존재할 경우는 4장 동적 객체를 추출하는 기법에서 이야기 할 것이다. 지역적 탐색은 식 (5)와 같다.

$$LD(x, y) = \sum_{y=1}^N \sum_{x=1}^N |I_a(x, y) - I_b(x, y)| \quad (5)$$

지역적 차 연산(Local Difference)은 기준 영상  $I_a$ 에서 참조 영상  $I_b$ 를 뺀 값으로 이를  $16 \times 16$  화소의 크기만큼 계산한다. 객체의 추출을 간단하게 하기 위해 직사각형으로 분리하는 방법과 정확한 추출을 위해 객체의 위치를 이용하여 정확히 뽑아내는 방법이 있다.

#### 4. 동적 모자이크 구축

##### 4.1 동적 객체 파라미터 측정

동적 모자이크 영상 생성을 위해서는 카메라의 움직임을 계산한 후에, 객체의 움직임을 인지해야 한다. 객체의 움직임을 위해서는 지역적 탐색(local search)이 이루어져야 한다. 지역적 탐색은 계산되어진 카메라의 움직임을 가지고서 두 영상간의 비교를 통해서 이루어 지게된다. 계산식은 식 (5)와 같다[15].

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

여기에서  $u(k)$ 는  $x$ 축,  $y$ 축으로의 방향 벡터이다. 비교되어지는 영상에서 방향벡터, 즉 카메라의 움직임을 거리만큼을 빼서 두 영상간의 에러 값을 계산한다. 이 때 두 영상사

이에서 객체의 움직임을 인지하는 임계 치를 크게 주면 커다란 객체의 움직임만 인식되고, 임계 치를 작게 주면 작은 움직임도 추출될 수 있다.

동적 객체가 존재하는 두 영상간에서는 카메라의 움직임과 객체의 움직임을 판단하는데 어려움이 있다. 카메라의 이동이 있다 할지라도 객체가 그 화면을 모두 채우고 있으면 카메라의 움직임을 인식하지 못하고, 카메라가 움직이지 않을 때도 커다란 객체가 움직이면 이를 카메라의 움직임으로 오인할 수 있다. 본 논문에서는 이런 모호성을 해결하기 위해 본 논문에서는 다음과 같은 방법을 제시한다. 가장 일반적으로 생각할 수 있는 간단한 방법은 카메라가 중요한 객체를 찍을 때는 주로 화면의 가운데를 차지하고 있다는 것이다. 그래서, 본 방법은 기준 영상의 가운데를 제외한 외곽 지역의 매크로 블록을 기준으로 유사성을 측정한다.

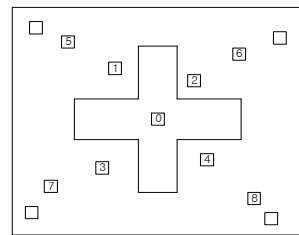


그림 3 기준영상의 매크로 블록

그림 3은 기준 영상의 비교할 매크로 블록을 보여준다. 각 매크로 블록은 십자가 모양의 중앙의 외곽 지역으로 비선형적으로 매크로 블록을 선택한다. 동적 객체를 판단하는 알고리즘은 다음과 같다.

1. 1,2,3,4에서 유사성을 측정하여 방향벡터 및 최소오류를 계산한다. 최소 값이 임계치 이상이면 제거한다.
2. 각 매크로 블록의 방향벡터가 모두 유사하면 그 방향벡터가 카메라의 움직임으로 인지한다.
3. 하나의 매크로 블록이라도 유사하지 않으면 6,7,8,9번의 매크로 블록으로 확장하여 검색한다.

4. 1번과 3번에서 계산되어진 방향벡터 중에 유사한 방향벡터를 선택한다.
5. 선택되어진 매크로 블록에서 affine 변환을 계산하여 카메라 파라미터를 측정한다.
6. 5번에서 계산된 카메라 파라미터를 이용하여 두 영상간의 차 영상을 식(4)를 이용하여 계산한다.

위의 알고리즘을 통해서 차 영상을 계산하여 차 영상이 임계치 이상이면 동적 객체가 존재한다고 판단한다.

#### 4.2 동적 객체 판별을 위한 영역 분할

동적 객체가 존재하는 경우에는 동적 객체 판별을 위한 영역 분할의 방법을 통하여 아래와 같은 방법으로 계산된다.

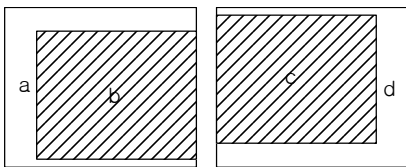


그림 4 동적 객체 판별을 위한 영역 분할

그림 4의 그림에서 b와 c의 빗금친 부분은 두 영상간의 유사 지역을 표시한다라고 가정한다. 위와 같은 가정을 통해 다음과 같은 결론을 유도할 수 있다.

1. a 영역의 어느 부분과 c 혹은 d 영역의 어느 부분이 일치하게 되면 이 지역은 객체가 되는 것이다. 2. b 영역과 d 영역의 부분이 같게 되면 이 부분 역시 동적 객체임을 알 수 있다. 마지막으로, 두 영상간에 일치하는 b 영역과 c 영역간에 일치하지 않는 부분은 객체임을 판별할 수 있다.

#### 5. 실험

본 논문에서 제안한 모자이크 시스템을 구현하기 위해서 HITACHI Digital Zoom 카메라로부터 영상을 입력받았다. 실험에 사용되어진 입력 영상은 320\*240의 크기의 컬러 영상으로 사용하였고, Pentium-450에서 Visual C++ 5.0을 이용하였다.

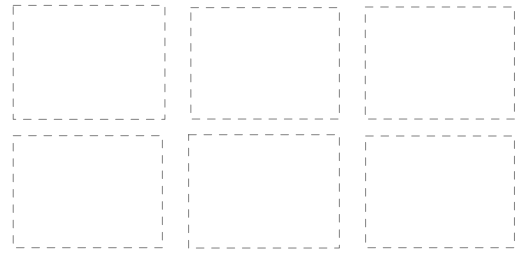


그림 5. 입력 영상들

그림 5는 동적 객체가 존재하지 않는 영상 중에 6개의 영상을 임의로 보여준 것이다.



그림 6. 정적 모자이크 영상

그림 6은 입력되어진 영상들을 정적 모자이크로 구축한 영상이다.



그림 7. 입력 영상들

그림 7은 동적 객체가 존재하는 영상들로서 임의의 4개의 영상을 보여준 것이다.



그림 8. 배경모자이크 영상

그림 8은 동적 객체가 존재하는 영상들로부터 배경 모자이크 영상을 구축한 영상이다.





## 그림 9. 동적 모자이크 영상

그림 9는 동적 객체를 일정한 시간별로 객체를 표현해 준 동적 모자이크로 구축되어진 영상이다.

### 6. 결론

본 논문에서는 모자이크 구축을 위해 이동하는 카메라를 통해 입력을 받아 모자이크 영상을 생성하는 방안을 제시하였다. 움직이는 카메라의 파라미터 측정 및 동적 객체의 파라미터를 측정하여, 배경과 객체간의 모호성을 방지하는 알고리즘을 제시하였으며, 동적 객체의 추출을 위한 영역 분할 방법을 제시하여 모자이크 영상을 정확하게 생성하였다.

또한, 많은 내용물을 포괄하고 있는 비디오로부터 영상들을 정합 하여 하나의 영상으로 표현해주는 모자이크 영상이 더욱 효과적으로 비디오의 정보를 나타내줄 수 있었다.

앞으로, 모자이크로 구축되어진 영상을 웹을 통하여 원 영상으로 재분할하는 방법에 대하여 연구할 것이며, 가상 세계(Virtual-Reality)에 모자이크 영상을 접목하여 학교 가이드 및 백화점 쇼핑몰 같은 응용 분야에 대하여 연구중이다.

### 7. 참고논문

- [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, P.Anandan, S.Hsu, "Mosaic based representation of video sequence and their applications", IEEE Proc ICCV'95, 605-611.

[5] Y.Xiong, K.Turkowski, "Creating image-based VR using a selfcalibrating fisheye lens", Proc.CVPR' 97, 237-243.

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

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

[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] Wonbo Yang, Moon sun Lim, Yangweon Rhee, "Video Mosaic System by Multi-Image", Korean Institute of Maritime Infomation & Communication Scienece , June, 1999, Vol.3 No.1, pp 104-108.

[13] Wonbo Yang, Jaewon Kim, Keunhwan Jeon, Yangweon Rhee, "Tree-based mosaic image and redistribution", Korea Information Processing Society, October 8, 1999