

효율적인 비디오 데이터베이스 구축을 위해 카메라와 객체 파라미터를 이용한 계층형 영상 모자이크

신성윤[†] · 이양원^{**}

요 약

모자이크 영상은 연관성 있는 비디오 프레임이나 정지 영상들을 합성하여 하나의 새로운 영상을 생성하는 것으로서 영상의 정렬, 통합 그리고 중복성에 대한 분석으로 수행된다. 본 논문에서는 효율적인 비디오 데이터베이스 구축을 위하여 카메라와 객체 파라미터를 이용한 계층형 영상 모자이크 시스템을 제시한다. 모자이크 영상 생성의 빠른 계산 시간과 정적 및 동적 영상 모자이크 구축을 위하여 트리 기반 영상 모자이크 시스템을 구축한다. 카메라 파라미터를 측정하기 위하여 최소사각형오류기법과 어파인 모델을 이용하였다. 객체의 추출을 위하여 차영상, 매크로 블록, 영역 분할과 4분할 탐색 기법들을 제시하여 사용하였다. 또한, 동적 객체 표현을 위하여 동적 궤도 표현 방법을 제시하였고, 완만한 모자이크 영상 생성을 위하여 블러링을 이용하였다.

A Hierarchical Image Mosaicing using Camera and Object Parameters for Efficient Video Database Construction

Seong-Yoon Shin[†] and Yang-Won Rhee^{**}

ABSTRACT

Image Mosaicing creates a new image by composing video frames or still images that are related, and performed by arrangement, composition and redundancy analysis of images. This paper proposes a hierarchical image mosaicing system using camera and object parameters for efficient video database construction. A tree-based image mosaicing has implemented for high-speed computation time and for construction of static and dynamic image mosaic. Camera parameters are measured by using least sum of squared difference and affine model. Dynamic object detection algorithm has proposed for extracting dynamic objects. For object extraction, difference image, macro block, region splitting and 4-split detection methods are proposed and used. Also, a dynamic positioning method is used for presenting dynamic objects and a blurring method is used for creating flexible mosaic image.

Key words: 영상 모자이크(image mosaic), 정적 및 동적 모자이크(static and dynamic mosaic), 비디오 데이터베이스(video database), 카메라와 객체 파라미터(camera and object parameter), 최소 사각형 오류 기법(least sum of squared difference), 어파인 모델(affine model), 차영상(difference image), 매크로 블록(macro block), 영역 분할(region splitting), 동적 궤도 표현(dynamic positioning)

1. 서 론

컴퓨터 정보 기술의 발달과 사용자의 욕구 증대로 인하여 멀티미디어 데이터(multimedia data) 처리에

관련된 연구가 폭넓게 진행되고 있다. 하지만, 멀티미디어 데이터의 방대한 양과 많은 처리 시간 때문에 이를 처리하고 전송하여 정보를 이해하는데까지는 실로 어려운 문제로 남아있다. 특히, 비디오(video)는 동영상 그 자체뿐만 아니라 텍스트(text), 정지영상, 사운드(sound) 그리고 보이스(voice) 등이 통합

[†] 정회원, 군산대학교 컴퓨터과학과 박사과정

^{**} 정회원, 군산대학교 컴퓨터과학과 교수

된 형태로서 많은 폭넓은 정보를 내포하고 있기 때문에 사용자가 비디오의 내용물을 빠르게 이해하는데 쉽지 않은 문제점을 가지고 있다. 또한, 하나의 비디오에서 장면과 장면의 내용들이 서로 다르므로 이들의 정보를 쉽게 표현해 줄 수 있는 방법의 구현이 쉽지 않으며 유사한 내용들을 지니고 있는 정지 영상의 경우 많은 저장공간을 차지한다는 문제점을 가지고 있다.

모자이크(mosaic) 영상이라는 것은 연관되어지는 여러 비디오 및 정지 영상들을 합성하여 하나의 새로운 영상을 생성하는 것을 말하며[1], 배경 중심의 정적 모자이크, 동적 객체 표현을 위한 동적 모자이크, 그리고 동적/정적 표현을 통합하여 대표적으로 나타내는 시놉시스(synopsis) 모자이크로 분류된다[2].

모자이크 구축 과정의 기술들은 연속된 영상의 정렬, 통합 그리고 중복성에 대한 분석으로 이루어진다. 파노라마 모자이크 시스템(panoramic mosaic system)에 대한 응용분야로 가장 전통적인 분야는 공중 인공위성 사진의 구축이며, 최근에는 장면 고정과 장면 감지, 비디오 압축과 색인, 그리고 카메라의 해상도 관련 연구는 물론, 심지어 간단한 사진 편집까지 매우 다양한 분야에서 연구되고 있다[2-4]. 비디오에서 하나의 장면을 대표하는 영상을 대표 영상이라 하는데, 대표 영상만을 가지고서 동영상의 한 장면을 이해하기에는 어려움이 많이 존재한다. 모자이크 영상은 유사한 많은 영상들을 하나의 영상으로 만들어주므로 동영상의 대표 영상에서 놓칠 수 있는 정보들의 모호성 문제점을 해결해준다. 모자이크 영상은 이웃하는 영상을 모두 포괄하고 있게 되므로 최소한의 저장 공간 사용, 빠른 자료의 전송, 그리고 전체적인 장면을 이해할 수 있다는 커다란 장점을 지니게 된다.

모자이크 시스템을 구축하기 위한 방법은 현재 많은 연구가 진행되어 왔으며, 커다란 이슈를 불러일으키며 더욱 발전하고 있는 실정이다. 많은 논문에서 차영상, 영상의 좌표 변환 기법, optical flow 그리고 움직임 측정에 대한 문제점과 해결 방법을 제시하였다.

Shum 등[3]은 일반적인 평면에 투영하는 방법이 아닌 각 입력 영상을 변환 행렬과 연관하여 모자이크 영상을 표현하는 방법을 제안하였다. 하지만 이 방법은 잡음에 매우 민감하다는 단점을 가지고 있다. Xiong 등[4]은 카메라를 90도로 회전하여 4개의 영

상을 얻어 가상세계를 구축하는 방안을 제시하였다. IBM사의 ImageMiner라는 시스템은 IBM사의 트래이드마크로서 동영상 분석에 정지 영상 분석을 통합한 시스템인데, 여기에서는 모자이크 영상을 분석하는 단계로서 색, 질감 그리고 윤곽 정보를 이용하여 객체를 인식하는 방법을 기술하였다[5]. 하지만 ImageMiner 시스템은 모자이크 영상을 색, 질감, 그리고 윤곽 정보만을 가지고 객체를 인식하였기 때문에 동적 객체는 인식하지 못하는 문제점을 가지고 있다. Irani 등[2]은 동적 객체에 대한 동적 모자이크 구축에 대하여 제안하였지만, 움직이는 카메라와 동적 객체를 판별할 수 있는 구체적인 방법은 기술하지 않았다.

모자이크 영상을 생성함에 있어 가장 쉬운 방법은 영상들 사이에서 단지 이동만 이루어졌다는 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[6-9].

모자이크를 구축하기 위해서는 반드시 카메라의 움직임을 측정해야 한다. 카메라 움직임 측정은 optical flow를 이용한 카메라의 파라미터(camera parameter) 계산을 통하여 수행하는 방법이 있다[3,10]. 또한, 카메라 파라미터를 추출하기 위해 사용하는 동작 모델에는 2차원 파라미터 동작 모델과 복잡한 3차원 동작 모델이 사용되고 있다[3]. 2차원 동작 모델 중에서 이동과 신축의 파라미터를 포함하면서 회전을 동시에 측정할 수 있는 어파인 모델(affine model)이 카메라 파라미터 측정에 가장 많이 이용된다[4].

본 논문에서는 정지 영상과 동영상을 입력받아 전처리 작업을 거친 후 카메라와 동적 객체의 파라미터를 측정하고 객체를 추출하여 모자이크 영상을 구축하는 방법을 제시하는데 전체적인 시스템 구조는 그림 1과 같다. 본 논문의 2장에서는 계층형 모자이크 영상 구축 방법에 대해 설명하고 3장에서는 정적 및 동적 모자이크 영상 구축에 대해 설명한다. 4장에서는 정적 및 동적 모자이크 영상 구축 실험을 수행하고 5장에서 결론을 맺고 향후 연구 방향에 대해 알아본다.

2. 계층형 모자이크 영상 구축

모자이크 영상 생성에서, 먼저 선행될 사항은 모

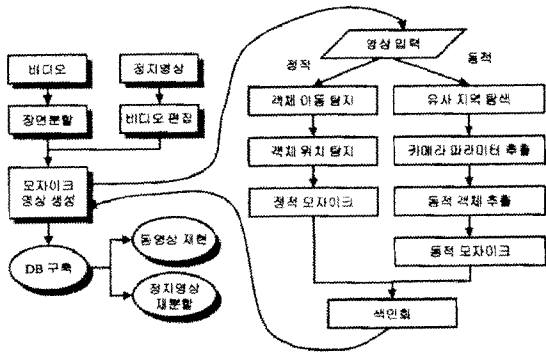


그림 1. 전체적인 시스템 구조

은 영상을 연속적으로 유사한 영상들끼리 정렬시키는 것으로, 이는 비교할 두 영상간의 일치하는 지역을 하나의 부분으로 정합 하는데 필요하다.

2.1 모자이크 시스템 흐름도

모자이크 시스템에서는 그림 2와 같이 비디오 카메라로 찍은 장면 속의 영상과 일반 카메라로 찍은 정렬된 영상이 입력으로 사용되는데, 일반 카메라로 찍혀진 영상은 각 영상들의 차이가 상당히 클 수 있다는 점이 있다. 이런 경우에 유사성 측정을 통하여 모자이크를 구축하는 것이 효과적이다.

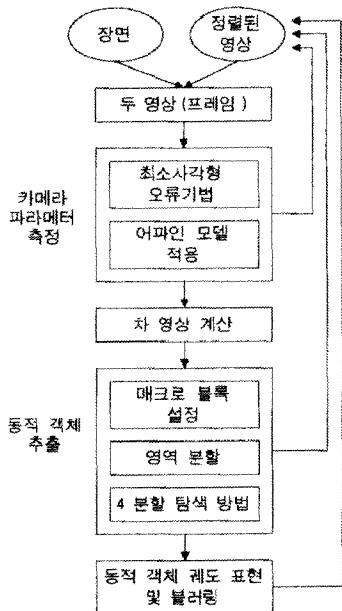


그림 2. 모자이크 영상 구축을 위한 흐름도

그림 2는 모자이크 영상 구축을 위한 흐름도를 나타낸다. 카메라로부터 얻어진 실세계의 동영상들은 분할되어진 장면으로써 연속되는 프레임(frame)으로 입력받고 정지 영상들은 자동 정렬된 영상을 받아 들여 두 영상 또는 프레임을 비교한다. 두 영상간의 카메라 이동을 측정하기 위해 최소사각형오류기법 (least sum of squared difference) 및 어파인 모델을 이용하였다. 동적 객체의 존재를 판단하기 위해 본 논문에서 제안하는 동적 객체 추출을 위한 매크로 블록(macro block) 설정, 영역 분할, 그리고 4분할 탐색 방법을 제시한다. 동적 객체의 존재 시에는 객체를 인지하여 추출한 후, 동적 객체의 레도 표현과 블러링(blurring) 기법을 적용하여 전체적으로 완전한 모자이크 영상을 구축한다.

2.2 트리(tree) 기반 모자이크

본 논문에서는 모자이크 영상 생성의 빠른 계산 시간을 얻기 위하여 트리 기반 모자이크 영상 생성 방법을 제시한다. 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.

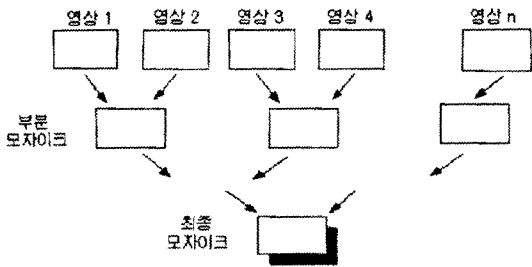


그림 3. 트리 기반 모자이크 영상

3. 정적 및 동적 모자이크 영상 구축

3.1 정적 모자이크 구축

정적 배경에서 동적 객체를 추출하는 방법에는

영상(difference image) 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 I_a 는 기준영상을, I_b 는 참조영상을 나타낸다. 또한 객체 부분을 탐지하기 위해서는 지역적인 탐색이 필요한데, 본 논문에서는 16×16 매크로 블록을 이용하여 객체를 탐지하도록 하며 식 (2)에서 표현한 것과 같이 지역적 차 연산을 수행하여 탐지한다.

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

지역적 차 연산은 기준영상 I_a 에서 참조영상 I_b 를 뺀 값으로 이를 16×16 화소의 크기만큼 계산한다. 이렇게 매크로 블록을 이용한 지역적 차 연산을 수행하여 객체 영역을 탐색한 후, 최소사각형오류기법을 통해서 두 영상의 유사성을 체크하고 카메라의 움직임 계산을 하여 각 세 점 이상의 좌표값을 가지고 어파인 파라미터를 추정하여 두 영상을 정합 함으로써 정적 모자이크 영상을 구축한다. 최소사각형오류기법과 어파인 파라미터 추정은 다음절의 동적 배경과 동적 객체를 이용한 모자이크 구축에서 자세히 다루도록 한다.

3.2 동적 모자이크 구축

동적 객체가 존재하는 두 영상간에서는 카메라의 움직임과 객체의 움직임을 판단하는데 어려움이 있다. 카메라의 이동이 있다할지라도 객체가 그 화면을 모두 채우고 있으면 카메라의 움직임을 인식하지 못하고, 카메라가 움직이지 않을 때에도 커다란 객체가 움직이면 이를 카메라의 움직임으로 오인할 수 있다.

본 논문에서는 이런 모호성을 해결하기 위해 다음과 같은 방법을 제시한다. 가장 일반적으로 생각할 수 있는 간단한 방법은 카메라가 중요한 객체를 찍을 때는 주로 화면의 가운데를 차지하고 있다는 것이다. 그래서, 본 방법에서는 기준 영상의 가운데를 제외한 외곽 지역의 매크로 블록을 기준으로 유사성을 측정한다.

그림 4는 기준 영상의 비교할 매크로 블록을 보여준다. 각 매크로 블록은 십자가 모양의 중앙의 외곽 지역으로서 비선형적으로 매크로 블록을 선택한다.

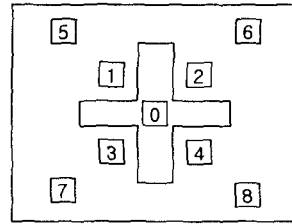


그림 4. 기준 영상의 매크로 블록

우선적인 가정은 객체의 크기는 영상의 반을 차지하지 않는다는 것이다. 이렇게 매크로 블록을 설정하여 동적 객체를 판단하는 알고리즘은 다음과 같다.

- (1) 두 영상(프레임)의 입력
- (2) 매크로 블록 1, 2, 3, 4의 유사성 측정:
if(최소 사각형 오류 측정 값 < 임계치)
then
 {방향 벡터 추정;
 goto (3);}
else
 goto (1);
- (3) if(방향 벡터가 4개의 매크로 블록과 모두 유사)
then
 {카메라 파라미터 = 방향 벡터;
 goto (4);}
else
 {매크로 블록 5, 6, 7, 8로 확장 검색으로
 방향벡터 추정;
 goto (4);}
- (4) 어파인 변환 계산을 통한 카메라 파라미터 측정
- (5) 두 영상의 지역적 차 연산

위의 알고리즘을 통해서 차 연산을 수행하여 결과가 임계치 이상이면 동적 객체가 존재한다고 판단한다. 다음으로 위의 알고리즘에 사용된 각각의 방법들에 대하여 알아본다.

3.2.1 최소사각형오류기법

먼저, 두 영상간 유사성을 찾기 위해 정확한 카메라 파라미터를 추출해야 하는데, 본 논문에서는 식 (3)과 같이 일정 윈도우 블록에 대한 최소사각형오류기법을 제시한다.

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

식 (3)에서 X는 x, y의 픽셀 위치를 말하고, b는 한 영상의 정사각형의 윈도우를 일컫는다. 기준영상 I_i에서 I_j의 영상의 차이 값의 제곱으로 계산되어지는 곳마다 참조영상 I_j의 모든 d_k의 블록을 통하여 최소 사각형 에러 값을 계산한다. 이 때 계산되어진 사각형 에러값 중 최소의 값을 선택한다.

3.2.2 어파인 모델

이렇게 식 (3)을 통해 가장 유사한 화소 값을 찾아내 이를 바탕으로 어파인 모델을 사용하여 영상의 회전, 신축 및 이동을 포함하는 카메라 파라미터를 측정한다. 어파인 모델은 식 (4)와 같다.

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

어파인 모델을 계산하기 위해서는 기준 영상과 참조 영상의 비선형의 세 개의 점들 (x₁, y₁, x₂, y₂, x₃, y₃) (x₁, y₁, x₂, y₂, x₃, y₃) 이 필요하다. 즉, 다음과 같은 선형방정식이 3개가 있어야 한다.

$$x' = ax + by + e \quad y' = cx + dy + f \quad (5)$$

이 세 개의 점들을 통하여 세 개의 선형방정식을 만들어 어파인 모델의 6개의 파라미터(a, b, c, d, e, f)를 계산한다. 기준 영상의 세 점과 참조 영상의 세 점을 통해서 세 개의 선형방정식을 계산하여 6개의 파라미터를 계산한 후 참조 영상의 모든 픽셀에 대응하는 점들을 배치한다. 식 (5)에서 a는 x축으로의 신축 계수, b는 x축으로의 회전 계수, c는 x축으로의 이동 계수를 말하고, d는 y축으로의 신축 계수, e는 y축으로의 회전 계수, f는 y축으로의 이동계수를 말한다.

3.2.3 지역적 차 연산

다음으로 객체의 움직임을 인지해야 하는데 이를 위해서는 지역적 탐색이 이루어져야 하는데, 본 논문에서는 제시하는 방법은 식 (6)과 같이, 계산되어진 카메라 움직임을 가지고 두 영상간의 방향 벡터와 매크로 블록을 이용한 차 연산을 통한 비교로 수행한다.

$$E(O) = \sum_{h \in H} [I_i(X+b) - I_j(X-u(k)+b+d_k)]^2 \quad (6)$$

여기에서 u(k)는 x축, y축으로의 방향 벡터로서

카메라의 움직임의 거리만큼을 나타내며 이를 빼서 두 영상간의 에러 값을 계산한다. 이때 객체의 움직임을 인지하는 임계치를 크게 주면 커다란 움직임만 인식되고, 작게 주면 작은 움직임도 추출할 수 있다.

3.2.4 영역 분할

또한, 본 논문에서는 동적 객체 판별을 위한 기본적인 가정을 통해 영역 분할 방법을 제시한다. 그림 5에서 두 영상의 b와 c의 부분은 두 영상간의 유사 지역을 표시한다라고 가정하면 동적 객체를 추출하기 위해서는 다음과 같은 가정이 따른다.

첫째, a 영역의 어느 부분과 c 혹은 d 영역의 어느 부분이 일치하게 되면 이 지역은 객체가 되는 것이다. 둘째, 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.

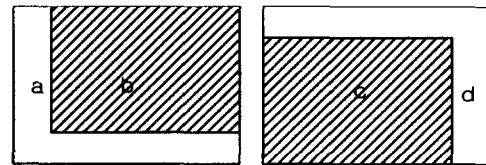


그림 5. 동적 객체 판별을 위한 영역분할

3.2.5 4분할 탐색

동적 객체의 추출은 두 영상 내에서 동적 객체가 존재하는 유사 지역을 탐색하여 동적 객체만 추출하는 것으로서 두 유사 지역의 차 영상을 계산한 값이 큰 경우에는 두 영상 내에 동적 객체가 존재한다고 가정한다. 그림 6에서 나타내는 것과 같이 1사분면의 지역을 계산하고, 값이 작으면 2사분면으로 계속적으로 진행해나간다. 동적 객체의 크기는 최소 7×7 픽셀 크기 이하의 객체는 무시하고, 다중의 동적 객체가 검출된 경우 이 중의 가장 큰 동적 객체만을 추출한다. 마지막 8×8 블록까지 탐색하여 영역의 검사를 마쳐 블록들의 합으로 동적 객체의 영역을 설정한다.

3.2.6 동적 궤도 표현과 블러링

배경영상 합성은 동적 객체 추출 후 나머지 배경 영상만을 합성한다. 이때, 추출된 동적 객체의 부분

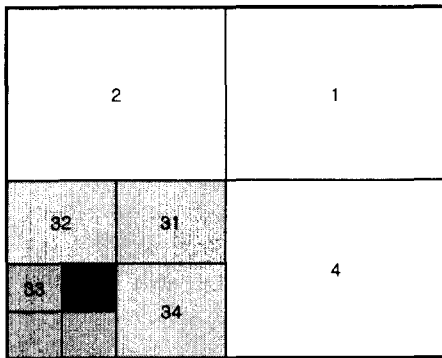


그림 6. 4 분할 탐색 방법

에는 나머지 다른 영상의 배경 부분이 삽입된다. 배경 모자이크 영상을 생성한 후 동적 객체를 표현할 경우 추출된 객체의 가로, 세로의 최대 크기의 1.5배 이상의 거리가 떨어진 동적 객체를 표현해 주도록 하는데 본 논문에서 제시한 평가 함수 식은 다음 식 (7)과 같다.

if ($A > 1.5B$) *then* 동적객체표현
 where $A = I_i(x2, y2) - I_{i-1}(x1, y1)$
 $B = \text{Length}(O_{\text{Max}}[(x1, x2), (y1, y2)])$ (7)

식 (7)은 동적 객체의 최대 크기(O_{Max})를 계산한 후, 동적 객체가 오른쪽으로 이동한 입력물인 경우 현재 영상의 왼쪽 부분과 이전 영상의 오른쪽 부분의 거리 차이 값($I_i(x2, y2) - I_{i-1}(x1, y1)$)이 객체의 최대 크기의 1.5배가 넘는 경우에 표현한다.

블러링은 각각의 영상이 합성되는 경계부분에 3 X 3 마스크를 이용하여 전체 마스크의 합이 1이 되도록 설정하는 가장 일반적인 방법을 사용하여 시각적으로 완만한 모자이크 영상을 생성하도록 한다.

이렇게 생성된 모자이크 영상은 2.2절에 설명한 트리 기반 모자이크 형태로 비디오 데이터베이스에 저장된다. 즉, 실제 비디오 데이터베이스에 저장되는 영상은 완전히 구축된 모자이크 영상이다. 이 모자이크 영상에는 실제 각각의 객체와 배경들이 모두 함축되어 저장되어 있으며 이를 표현하는 방법에 따라 부분적인 묘사와 생략이 가능하다. 또한 각각의 모든 영상들을 저장하지 않으므로 저장 공간을 보다 효율적으로 사용할 수 있으며 모자이크 영상 하나에 대한 부분적인 검색들이 가능하므로 빠르게 검색할 수 있다.

4. 실험 및 결과

본 논문에서 제안한 모자이크 시스템을 구현하기 위해서 HITACHI Digital Zoom 카메라로부터 영상을 입력받아서 320 * 240 크기로 정규화 한 컬러 영상을 사용하였고 Pentium-4 1.3GHZ에서 Visual C++ 6.0을 이용하였다.

많은 실험 데이터 중에서 그림 7은 정적 모자이크를 구축하기 위해 학교 내부의 건물과 운동장을 찍은 18초의 동영상에서 전체 프레임의 일부를 디스플레이 해 준 것이다.

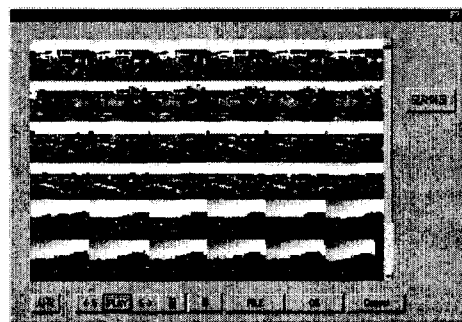


그림 7. 각 프레임 영상

그림 8은 그림 7의 전체 프레임들 중에서 임의의 6개의 영상만을 선택하여 나타내 준 것으로서 모자이크 영상을 생성하기 위한 입력에 사용된 영상이다.

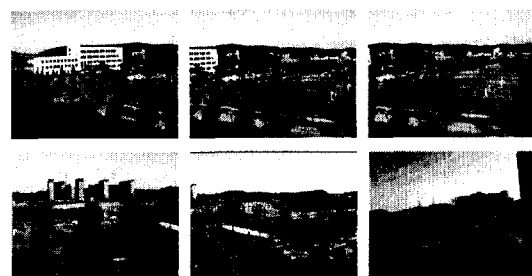


그림 8. 임의의 입력 영상들

그림 9는 그림 8에 나타난 6개의 입력받은 영상들로부터 생성되어진 정적 모자이크 영상이다. 동적 객체가 존재하지 않는 파노라믹 모자이크 영상으로 좌우가 상당히 길어진 것을 알 수 있으며, 이러한 모자이크 영상을 통해서 전체적인 동영상의 내용을 쉽게 이해할 수 있다.



그림 9. 정적 모자이크 영상

그림 10은 연구실 내부를 카메라로부터 찍은 26초의 동영상으로써 360도 회전한 동영상이며, 다음 그림 11은 입력으로 사용될 임의의 6개의 영상을 나타내 준 것이다.

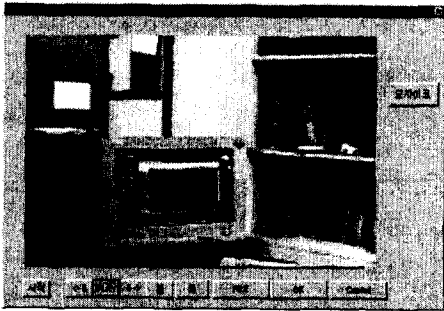


그림 10. 입력 동영상



그림 11. 임의의 입력 영상

그림 12는 그림 11의 360도 회전한 연구실 내부에서 임의로 이용한 6개의 입력 영상으로 생성한 정적 모자이크 영상이다. 이러한 360도 모자이크 영상을 통해 연구실 내부를 360도로 회전한 것처럼 이해할 수 있다.

다음 그림 13은 동적 모자이크 영상을 구축하기



그림 12. 360도 회전한 정적 모자이크 영상

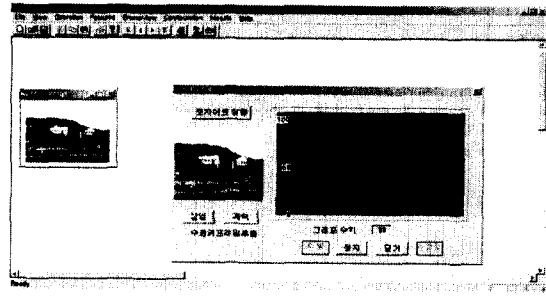


그림 13. 동적 모자이크 구축을 위한 전체화면

위한 전체 화면으로, 입력 동영상은 총 57초로 자동차의 동적 객체가 존재하는 영상인데, 동적 객체인 자동차가 존재하는 장면은 15초간이다.

다음 그림 14는 동적 객체가 존재하는 영상에서 동적 객체를 제거한 후의 배경 모자이크를 생성한 영상으로 동적 객체가 존재했던 위치에는 또 다른 프레임의 배경 영상으로 채움으로써 제거된 동적 객체의 부분 지역을 채운 것이다.



그림 14. 동적 객체를 제거한 배경 모자이크 영상 생성

그림 15는 추출된 동적 객체 영상을 그림 14의 배경 모자이크를 구축한 영상 위에 표현해 준 것이다.



그림 15. 객체를 표현한 동적 모자이크 구축 영상

그림 16은 실내에서 찍은 연구실 내의 동적 객체가 존재하는 동영상으로써 임의의 4개의 영상을 보여준 것이다.

그림 17은 동적 객체를 일정한 시간별로 객체를 표현해 준 동적 모자이크로 구축되어진 영상이다. 여



그림 16. 임의의 동적 입력 영상들

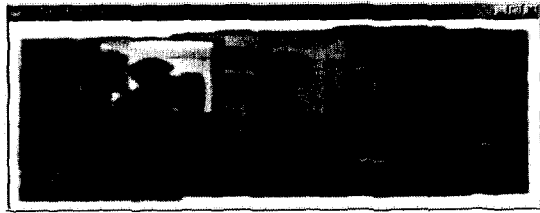


그림 17. 동적 모자이크 영상

기에서 동적 객체를 표현해 준 것은 실험에 의해 80 픽셀 이상 이동되었을 때 나타내주었다.

교내에서 촬영한 30초 짜리 동영상 5개에서 각각 10개씩의 프레임을 추출하여 모자이크 영상을 구축한 결과 다음 표 1과 같이 저장 공간에서 35.1%라는 커다란 절감률을 보였다.

표 1. 입력영상에 대한 저장공간 절감률

입력영상 (프레임) 수	평균 영상크기	전체영상 저장크기	모자이크 영상 저장크기	저장공간 절감률
10개	75K	750K	487K	35.1%

이러한 절감률은 보다 효율적인 비디오 데이터베이스를 구축하는데 커다란 도움을 줄뿐만 아니라 전체적인 내용의 파악과 이해를 돕는데도 큰 몫을 차지한다.

거의 대부분의 모자이크 시스템에서 시스템의 환경이나 시스템 성능, 그리고 장면전환 검출이나 정규화 작업과 같은 다양한 전처리 과정들에 따른 제약조건들 때문에 보다 구체적이고 계량적인 성능 측정 기준과 결과의 제시를 지양하고 있으므로 다른 시스템들과의 비교 평가는 사실 어려운 실정이다. 하지만 본 논문에서는 기존의 모자이크 시스템과 비교하여 본 연구 및 실험의 독창적이고 두드러진 특징 요소들을 표 2와 같이 요약할 수 있다.

5. 결 론

본 논문에서는 효율적인 비디오 데이터베이스 구

표 2. 기존 시스템과의 특징 요소 비교

구분	기존 시스템	본 시스템
구축 기반	연속적 프레임 기반	트리 기반
적용 분야	정적 또는 동적	정적과 동적 모두
연산 수행 영역	전체 영역	부분 영역
측정 파라미터	주로 카메라 파라미터	카메라와 객체 파라미터 모두
동적 객체 궤도 표현	불분명하며 자유롭지 못함	자유로움
모자이크 경계면	부드럽지 못함	완만함

축을 위하여 카메라와 객체 파라미터를 이용한 계층형 영상 모자이크를 구축하는 방안을 제시하였다. 최소사각형오류기법과 어파인 모델을 이용하여 카메라 파라미터를 측정하였고 두 입력 영상의 유사성 측정을 위해 차영상을 이용하였다. 또한 동적 객체의 추출을 위해 매크로 블록을 설정하여 동적 객체를 탐색하고 영역 분할과 4분할 탐색 방법을 이용하여 객체를 추출하였다. 그리고 동적 객체의 표현을 위해 동적 궤도 평가 함수를 이용한 후 블러링을 통하여 완만한 모자이크 영상을 구축하였다.

전체적으로 트리 기반 계층형 모자이크 영상의 구축은 물론, 정적 및 동적 모자이크 영상의 구축을 통하여 보다 효율적인 비디오 데이터베이스를 구축하는데 있어서 빠른 시간과 저장공간을 절감할 수 있는 기반을 제시하였다.

향후 연구 과제로서 객체의 움직임을 보다 정확하게 판단할 수 있는 방법의 개발과 웹 상에서 제약을 받지 않는 모자이크 영상 구축 방법에 대한 연구가 필요하다고 사료된다.

참 고 문 헌

[1] Wonbo Y., Jaewon K., Keunhwan J., and Yangweon R., "Tree-Based Mosaic Image and Redistribution," *Proc. KIPS*, Oct. 8, 1999.
 [2] M. Irani, P. Anandan, and S. Hsu, "Mosaic Based Representation of Video Sequence and Their Applications," *IEEE Proc. ICCV'95*, pp. 605-611, 1995.
 [3] Heung-Yeung Shum and Richard Szeliski,

"Panoramic Image Mosaic," *MSR-TR-97-23*, 1997.

[4] Y. Xiong and K. Turkowski, "Creating Image-Based VR Using A Selfcalibrating Fisheye Lens," *Proc. CVPR' 97*, pp. 237-243, 1997.

[5] J. Krey B, M. Roper, P. Alshuth, Th. Hermes, and O. Herzog, "Video Retrieval by Still-Image Analysis with ImageMiner," *Storage and Retrieval for Image and Video Databases (SPIE)*, pp. 36-44, 1997.

[6] A. Tekalp, M. Ozkan, and M. Sezan, "High-Resolution Image Reconstruction from Lower-Resolution Image Sequences and Space-Varying Image Restoration," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. III-169-172, March 23-26, 1992.

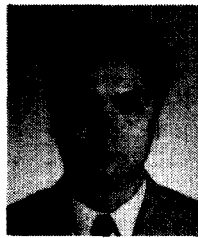
[7] L. Teodosio and W. Bender, "Salient Video Stills : Content and Context Preserved," *Proc. ACM Multimedia Conf.*, 1993.

[8] Gulrukh Ahanger and Thomas D.C. Little, "A Survey of Technologies for Parsing and Indexing Digital Video", *Journal of Visual Communication and Image Representation*, Vol. 7, No.1, March, pp.28-43, 1996.

[9] D. J. Heeger, "Optical Flow Using spatio-temporal Filters", *International Journal of*

Computer Vision, pp. 279-302, 1998.

[10] Richard Szelisk, and H. Shum, "Creating Full View Panoramic Image Mosaics and Environment Maps," *In Proc. of SIGGRAPH*, pp. 251-258, 1997.

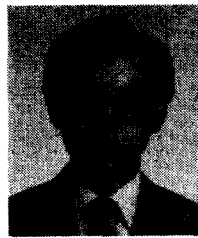


신 성 운

1994년 군산대학교 컴퓨터과학과 (이학사)
 1997년 군산대학교 정보통신공학과(공학석사)
 2000년~현재 군산대학교 컴퓨터과학과 박사과정
 1998년~2002년 군장대학 컴퓨터

응용학부 겸임교수

1999년~현재 (주)네트플러스 동영상연구팀 연구원
 관심분야 : 멀티미디어, 컴퓨터 비전, 인공지능



이 양 원

1978년 숭실대학교 전자계산학과 (공학사)
 1983년 연세대학교 전자계산학과 (공학석사)
 1994년 숭실대학교 전자계산학과 (공학박사)
 1979년~1986년 한국국방연구원

정보관리위원회 연구원

1986년~현재 군산대학교 컴퓨터과학과 교수
 관심분야 : 멀티미디어, 컴퓨터비전, 인공지능, 가상현실