

논문 2010-1-13

동심원 모자이크의 실시간 표현을 위한 데이터 압축 알고리즘에 관한 연구

A Study on the Data Compression Algorithm for Just-in-Time Rendering of Concentric Mosaic

지인호*, 안홍영**

Inn-Ho Jee*, Hong-Yeoung Ahn**

요약 동심원 모자이크는 공통의 공간적인 기준을 사용하여 비디오 분할의 프레임들을 정렬시키거나 합쳐서 형성된다. 과거의 3차원 웨이브렛 변환 코딩의 연구와 비교하여 유연한 부분적 부호화와 비트 스트림의 임의 검색을 가능하게 하여주는 중요한 구성을 고려하여 압축된 동심 모자이크의 just-in-time(JIT) 표현 알고리즘을 제안하였다. 이 동심 모자이크의 JIT 표현이 실시간 표현을 수행하기 위해서는 많은 계산량이 요구된다. 압축된 형태로 데이터가 유지되어 특정 장면을 표현하기 위해 사용되는 내용만이 복호 되어야 한다. 그래서 제안된 알고리즘은 웨이브렛을 리프팅 계산식의 구조를 사용하여 동심 모자이크의 실시간 표현을 가능하게 해준다.

Abstract Concentric mosaics are made with arranging and summing of video frames by using common spacial standards. Compared with previous works on 3-D wavelet transform coding, we have made important design considerations to enable flexible partial decoding and bit-stream random access. A just-in-time(JIT) rendering engine of the compressed concentric mosaic is developed. However, computationally, it is still demanding to accomplish the real-time rendering. Only the contents for specific scene representation are need to be decoded by maintaining compressed data. Thus our proposed algorithm is able to render real concentric mosaic by using lifting scheme instead of wavelet transform.

Key Words : Concentric Mosaic, Lifting, SPIHT, EZW

1. 서론

동심원 모자이크는 공통의 공간적인 기준을 사용하여 비디오 분할의 프레임들을 정렬시키거나 합쳐서 생성된다. 동심원 모자이크는 평면 원내에서 관찰자의 운동을 제안시키는 3차원의 plenoptic 함수를 사용한다. 동심 모자이크의 실제적인 환경은 카메라를 밖으로 향하게 하고 빛이 회전하는 것처럼 주사함으로서 빛의 끝에서 카메라를 돌리면서 얻어진다. 동심 모자이크의 데이터의 양은

엄청나므로 압축은 필수적이다. 동심원 모자이크의 교차 프레임과 프레임내의 높은 상관도를 이용하면 고압축을 얻을 수 있다^[1]. 3차원 웨이브렛 변환은 효과적인 비상관도, 에너지 포장도, 공간/시간 구역성, 다해상도 등의 바람직한 특성을 가지고 있다. 동심 모자이크의 데이터 구조는 천천히 움직이는 파노라마 카메라와 관련된 비디오 순차로서 정지화상과 비디오 압축기술이 적용되어야 한다. 그러나 동심의 모자이크는 새로운 압축기술이 필요한 특징이 있다.

첫째로 동심 모자이크는 영상들 사이에 잘 정비된 카메라의 이동과 관련된 1차원 영상배열이다. 교차 프레임

*정희원, 홍익대학교 컴퓨터정보통신공학과
**정희원, 홍익대학교 컴퓨터정보통신공학과
접수일자 2010.2.1, 수정일자 2010.2.16

상관도가 보통의 비디오 신호보다 크다. 둘째로 동심 모자이크의 왜곡 허용한계는 작다. 왜냐하면 각각의 묘사된 동심원 모자이크가 정적으로 보여지고 인간 시각시스템이 시변 왜곡보다 정적 왜곡에 보다 민감하게 작동되기 때문이다. 사실 압축된 동심 모자이크는 데이터가 많아서 오늘날 대부분의 하드웨어가 그것을 처리하는데 어렵게 된다. 압축된 형태로 데이터가 유지되어 특정 장면을 표현하기 위해 사용되는 내용만이 복호 되어야 한다. 현재 관찰을 묘사하는 것이 필요하다. 이것을 just-in-time(JIT) 묘사라고 불러진다^[2]. JIT는 영상기반의 묘사압축과 묘사 알고리즘에 매우 중요한 요구사항이 된다.

II. 동심원 모자이크

카메라 회전으로 일정 주기의 영상을 촬영하고 촬영된 사진은 모자이크 영상으로 재배열하면 동심 모자이크가 생성된다. 카메라가 회전하는 동안 취해진 촬영을 $c(n, h, v)$ 표현하면, n 은 카메라 shot를 나타내고 h 는 shot 내에서 수평적인 위치를 나타내고, v 는 수직적인 위치를 나타낸다. 동심 모자이크 배경에서 촬영된 shot들이 모자이크 영상으로 재배열되는데, 모자이크 $F(h) = \{c(n, h, v) | n, v\}$ 는 모든 카메라 shot 위치인 수평위치 h 에서 수직적 slit으로 구성되어있다. 다른 배경에서 3개가 결합된 모자이크가 그림 1에 나타내었다^[3].

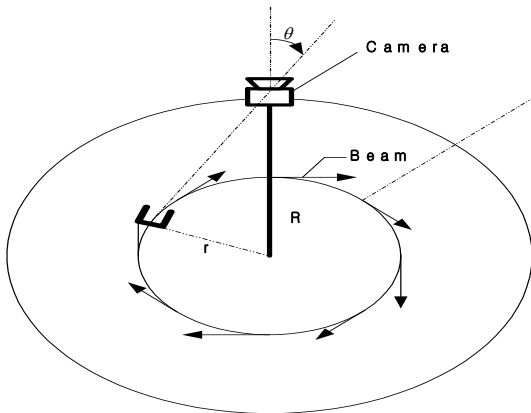


그림 4. 동심원 모자이크의 기하학적 표현
Fig. 1. Concentric mosaic imaging geometry

1. 동심원 모자이크의 정렬

모든 두 개의 인접한 모자이크 영상들은 주기적으로 정합과 정렬을 하여서 그 결과 모자이크들 간의 상관도는 최대가 된다. 두 개의 주기적으로 이동된 모자이크 사이의 평균 절대 에러(MAE)는 다음과 같이 계산된다.

$$MAE(w) = MAE[F(w), F(w-1)] \\ = \frac{1}{HV} \sum_{v=1}^V \sum_{h=1}^H |c(n-\Delta(w), w, h) - c(n-\Delta(w-1), w-1, h)| \quad (1)$$

여기서 H 와 V 는 각각 카메라 촬영 시 수직과 수평의 해상도를 나타낸다. 그림 2는 촬영된 각각의 영상으로부터 식(1)을 사용하여 모자이크 이미지를 얻는 과정을 보여주고 있다.

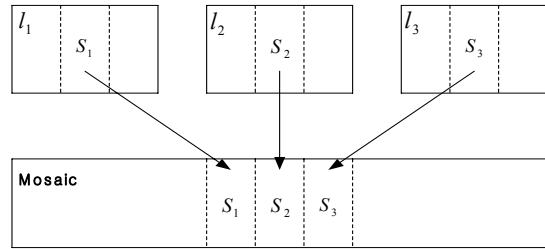


그림 5. Mosaic 영상의 정렬
Fig. 2. Alignment of Mosaic image

III. 리프팅을 이용한 3차원 압축

1. 리프팅 변환

리프팅 변환^[4]은 2세대 웨이브렛이라고 불리워진다. 리프팅 변환은 효과적인 비상관도, 에너지 포장도, 공간/시간 구역성, 다해상도 등의 바람직한 특성을 가지고 있다. 그림 3은 1차원 9/7리프팅 구조를 표현한 것이다. 기존의 컨벌루션 연산은 8번의 덧셈과 4.5번의 곱셈 과정을 통해 변환 계수를 얻지만, 9/7 리프팅 필터의 평균 연산은 계수 당 4번의 덧셈과 2번의 곱셈을 수행하므로 연산량이 거의 절반이다. 또한 하나의 계수를 얻기 위해 6개의 메모리 공간을 필요로 하는 in-place 연산 또한 가능하다.

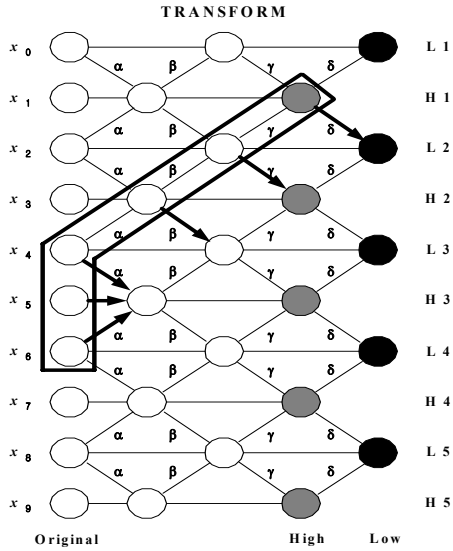


그림 6. 1차원 순방향 리프팅 변환
Fig. 3. One-dimensional forward lifting transform

2. 3차원 리프팅 변환

그림 4는 1-level의 3차원 리프팅을 보여주고 있다. Frame 사이의 리프팅을 거친 후 수평, 수직 방향으로 리프팅 과정을 거치게 된다.

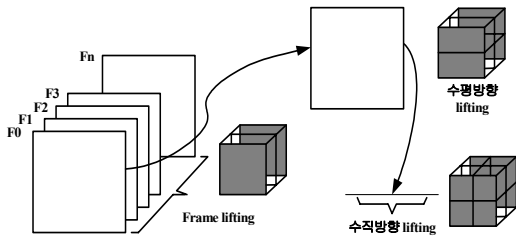
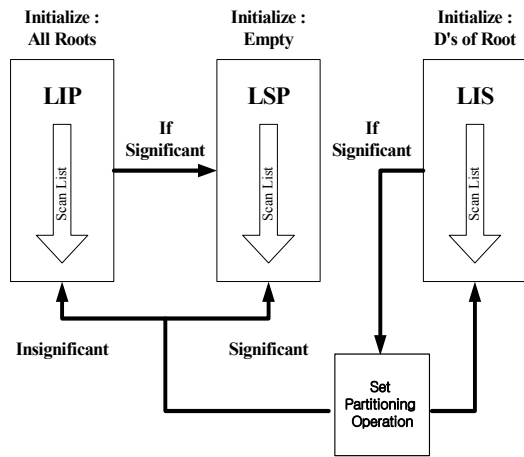


그림 7. 3차원 리프팅 변환
Fig. 4. 3-dimensional lifting transform

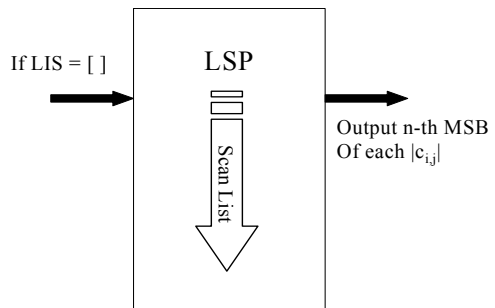
3. Set Partitioning In Hierarchical Trees

1996년 Said와 Pearlman^[5]은 웨이브렛 영역에서 웨이브렛 계수들의 부대역간 상관관계를 이용한 SPIHT(Set Partitioning in Hierarchical Trees)를 제안하였다. 웨이브렛 변환을 통해 다해상도로 분해된 부대역들을 얻을 수 있다. 각 부대역 계층에서 상위계층에 에너지가 집중되며, 하위 계층으로 갈수록 분산이 감소하게 된다. SPIHT는 EZW방식을 개선한 알고리즘으로 압축 효율을 개선한 알고리즘이다. SPIHT는 계산과정이 빠르고, 부

분 정렬에 의해 중요 비트를 먼저 보냄으로써 임베디드 부호화를 가능하게 하며, 점진적 전송이 가능하여 전송 과정에서의 오류에도 복호화 할 수 있는 장점을 가지고 있다. 또한 부대역간의 유사성을 이용하여 부호화함으로써 압축 효율을 높이고 원하는 비트율까지 부호화 할 수 있는 장점이 있다. EZW는 임베디드 부호화 과정을 모두 마친 데이터들은 단단계 적응적 산술부호화를 이용한 엔트로피 코딩이 필수적이다. 그러나 SPIHT의 경우 하위 집합 분할은 매우 효율적이고 임베디드 부호화 과정을 모두 마친 데이터들은 매우 간결해서 엔트로피 코딩을 거치지 않은 이진 전송만으로도 성능 면에서 대략 같거나 우수한 성능을 얻을 수 있다.



(a) SPIHT의 정렬 과정
(a) Sorting pass of SPIHT



(b) 중요 계수의 세분화 과정
(b) Refinement pass of LSP

그림 5. SPIHT의 부호화 알고리즘
Fig. 5. Encoding Algorithm of SPIHT

IV. Just - In - Time

동심 모자이크의 데이터 구조는 천천히 움직이는 파노라마 카메라와 연관된 비디오 순차로서 정지화상과 비디오 압축기술이 적용되어야 한다. 그러나 동심의 모자이크는 새로운 압축기술이 필요한 특징이 있다. 사실 압축된 동심 모자이크는 데이터가 많아서 오늘날 대부분의 하드웨어가 그것을 처리하는데 어렵게 된다.

압축된 형태로 데이터가 유지되어 특정 장면을 표현하기 위해 사용되는 내용만이 복호 되어야 한다. 현재 관찰을 묘사하는 것이 필요하다. 이것을 just-in-time(JIT) 묘사라고 불러진다. JIT는 영상기반의 묘사압축과 묘사 알고리즘에 매우 중요한 요구사항이 된다.

이러한 알고리즘을 만족하기 위해 먼저 3-D 리프팅 변환을 실행한다. 그런 다음 변환된 데이터 집합은 각 프레임 단위로 SPIHT 압축 알고리즘이 적용된다. 데이터가 압축되어 질 때 압축 비트율을 결정하는 것은 매우 중요하다. 각각의 프레임들은 서로 다른 비트율로 압축되어 진다. 저주파수 대역에는 높은 비트율을 제공하고 고주파수로 갈수록 낮은 비트율을 제공하게 된다. 각각의 프레임의 비트율들은 각 프레임들 에너지를 계산해 그 비율로 비트율을 계산할 수 있다. 만약 각 프레임들이 동일한 비트율을 갖게 된다면 복원된 프레임들의 성능은 저하되게 된다. 이렇게 임의의 비트율로 압축된 데이터에서 특정 장면을 추출하기 위해서는 먼저 전송된 비트열에서 헤더 정보를 읽어 각각의 프레임의 압축 비트율을 알고 있어야 한다. 또한 기준 프레임을 선정하여야 하는데 본 논문에서는 16프레임 단위로 압축을 실행하여 첫 번째 프레임을 기준 프레임으로 선정하였다.

그림 6은 9/7 필터를 사용한 1차원 역방향 리프팅 변환을 보여주고 있다. 먼저 기준 프레임이 될 첫 번째 프레임을 비트열에서 복원해야 한다. 그림 6에서 보는 것과 같이 \hat{x}_0 가 복원되기 위해서는 4개의 계수가 필요로 하게 되는데 두개는 저주파 대역 계수와 두개의 고주파 대역 계수가 된다. 고주파 대역의 계수는 가장 고주파수 성분이 되고 저주파 대역의 계수는 분할 된 계수들을 합성하여 얻어지게 된다. 이 과정을 거쳐 기준이 될 첫 번째 프레임을 복원하게 되면 다음 프레임의 복원은 첫 번째 프레임을 복원하는 과정에서 버퍼에 저장된 값을 이용하여 보다 쉽게 얻어질 수 있다.

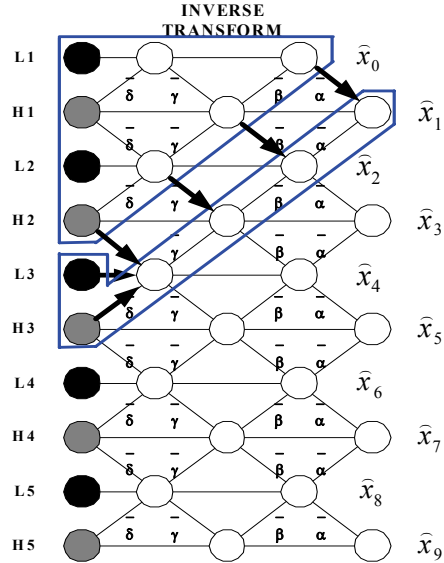


그림 6. 1차원 역방향 리프팅 변환
Fig. 6. One-dimensional reverse lifting transform

두 번째 프레임을 얻기 위해서는 그림 6에서 보는 것과 같이 기준 프레임을 얻는 과정에서 얻어지는 버퍼에 저장된 4개의 계수값과 새로이 추가되는 저주파 계수 1개 고주파 계수 2개 등 총 6개의 계수를 이용하여 다음 프레임의 계수들을 얻을 수 있다. 또한 이렇게 하여 \hat{x}_1 를 얻는 과정에서 \hat{x}_2 또한 동시에 얻을 수 있다. 따라서 홀수 프레임을 연산 하는 과정에서 짝수 프레임은 부가적으로 얻을 수 있는 장점이 있다.

V. 실험 및 결과

실험은 352x240 크기의 8 bit gray-level의 tennis 영상 16 프레임과 영상 512x512 크기의 8 bit gray-level의 실험영상 16 프레임을 사용하여 실험을 진행 하였다. 제안한 그림 6의 3-D 리프팅 구조로 웨이브렛 변환을 대체하여 복원하였다. 그림 7은 0.3 bpp로 압축된 비트열에서 얻은 처음과 중간인 1, 2, 8, 9 프레임의 tennis 영상을 나타내었다. 복원한 회복 성능은 다음의 PSNR공식으로 측정하였다.

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right)$$

리프팅 변환은 시간과 공간 영역에서 수행하였다. 표 1은 두 가지 실험 영상들의 0.3 bpp와 1.0 bpp에서의 압축 성능을 보여주고 있다. 실험 영상과 tennis 영상의 성능 차이가 나는 것은 실제로 연구실에서 촬영한 실험영상의 처리에서 발생한 결과로 판단된다.

표 1. 실험결과
Table 1. Experimental Results

압축률	PSNR(dB)	
	tennis	실험영상
0.3bpp	31.15	22.29
1.0bpp	36.34	26.90

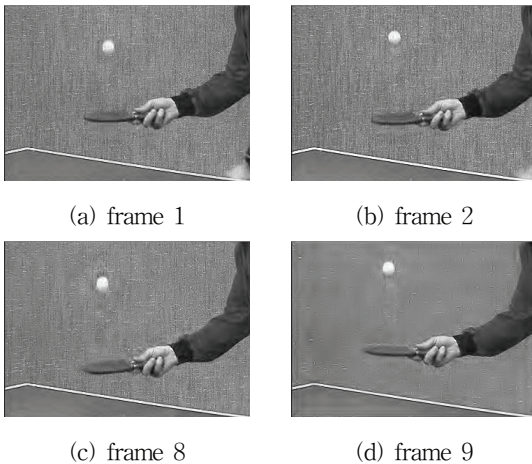


그림 7. 0.3 bpp로 압축된 비트열에서 얻은 영상
Fig. 7. Detected video sequence at 0.3bpp bit-stream



그림 8. 1.0 bpp로 압축된 실험영상에서 검색된 영상을 이용하여 생성된 모자이크 영상
Fig 8. Mosaic image using detected sequence at 1.0 bpp bit-stream

VI. 결론

동심원 모자이크의 데이터 구조는 천천히 움직이는 파노라마 카메라와 연관된 비디오 순차로서 정지화상과 비디오 압축기술이 적용되어야 한다. 그러나 동심의 모자이크는 새로운 압축기술이 필요한 특징이 있다. 또한 압축된 데이터라 하더라도 데이터의 양이 엄청나기 때문에 압축된 형태로 데이터가 유지되어 특정 장면을 표현하기 위해 사용되는 내용만이 복호 되어야 한다. 현재 관찰을 묘사하는 것이 필요하다. 본 논문에서는 3차원 리프팅 변환된 영상을 2차원 압축과정을 거쳐 동심원 모자이크의 실시간 표현을 가능하게 하기 위한 알고리즘을 제안하였다.

참고 문헌

- [1] H. Y. Shum and L. W. He, "Rendering with concentric mosaics," in Computer Graphics Proc., Annu. Conf. Series(SIGGRAPH'99) Los Angeles, CA, Aug. 1999, pp. 299-306.
- [2] Lin Luo, Yunnan Wu, Jin Li, Ya-Qin Zhang, "3-D Wavelet Compression and Progressive Inverse Wavelet Synthesis Rendering of Concentric Mosaic", IEEE Trans. Image Processing, vol. 11, pp. 802- 816, July 2002.
- [3] Sing Bing Kang , Yin Li , Xin Tong , Heung-Yeung Shum, Image-based rendering, Foundations and Trends® in Computer Graphics and Vision, v.2 n.3, p.173-258, January 2006.
- [4] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," J. Fourier Anal. Applicat., vol. 4, pp. 247-269, 1998.
- [5] A Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," IEEE Trans. Circuits Sys. Video Technol., vol. 6, pp. 243-250, June 1996.
- [6] B. J. Kim and W. A. Pearlman, "An embedded wavelet video coder using three-dimensional set partitioning in Hierarchical trees(SPIHT)," Proc.

IEEE Data compression conf, pp.251-260, Mar 1997.

- [7] Chi-Wing Fu , Liang Wan , Tien-Tsin Wong , Chi-Sing Leung, The rhombic dodecahedron map: an efficient scheme for encoding panoramic video, IEEE Transactions on Multimedia, v.11 n.4, p.634-644, June 2009.

저자 소개

안 홍 영(정회원)



- 1975년 : 서울대학교 전자공학과 공학사
- 1986년 : (미) University of Florida, Dept. of Electrical and Computer Eng. (MS)
- 1991년 : (미) University of Florida, Dept of Electrical and Computer

Eng. (Ph.D)

- 1975년~1984년 : 국방과학연구소(ADD) 선임 연구원
 - 1984년~1991년 : Research Assistant, University of Florida, USA
 - 1992년~현재 : 홍익대학교 컴퓨터정보통신공학과 부교수
- <주관심분야 : Wireless Networks, Ad-hoc Networks, Sensor Networks, Wireless Internet, VoIP, 4세대 이동통신, BeN, 영상신호처리>

지 인 호(정회원)



- 1980년 : 서울대학교 전자공학과 공학사
- 1983년 : 서울대학교 전자공학과 공학석사
- 1995년 : Polytechnic Institute of New York University, 전기 및 컴퓨터공학과 공학박사

- 1982년~1988년 : 국방과학연구소 선임연구원
 - 2004년~2005년 : University of Maryland at College Park, USA, 연구교수
 - 1995년~현재 : 홍익대학교 컴퓨터정보통신공학과 정교수
- <주관심분야 : CDMA/OFDM, 3D Image Processing, Multimedia Security, Multimedia Signal Processing>