

3D 디지털 영상의 DIBR 공격을 위한 강인성 블라인드 워터마킹

*이용석, **서영호, **김동욱
광운대학교

stucom12@kw.ac.kr, yhseo@kw.ac.kr, dwkim@kw.ac.kr

Robust and Blind Watermarking for DIBR attacked 3D Digital Image

*Yong-Seok Lee, **Young-Ho Seo, **Dong-Wook Kim
Kwangwoon University

요 약

본 논문에서는 3D 디지털 영상 콘텐츠의 저작권 보호를 위해 DWT(Discrete Wavelet Transform)를 이용하여 강인성(robustness), 비가시성(invisibility), 보안성(security)을 만족하는 블라인드(Blind) 워터마킹 방법을 제안한다. 이 방법은 깊이 영상 기반 렌더링(DIBR, Depth Image Based Rendering)에 효과적으로 대처하기 위해 고안된 알고리즘으로, 렌더링시 발생하는 왜곡에 영향이 적은 1-Level DWT LH 부대역에 워터마크를 삽입하는 방법이다. 이 방법에서 저작권에 대한 판단은 추출한 워터마크 정보를 통합 판단하는 방법을 사용한다. 이 방법은 저작권 보호를 위해 원본 영상을 보관할 필요가 없으며, 깊이 영상 기반 렌더링 baseline distance 5% 공격에 좌,우시점 영상 모두 효율적으로 적은 에러율을 보이며 워터마크로서 충분한 강인성을 보여준다.

1. 서론

최근 미디어 기술의 발달로 VR(Virtual Reality)을 쉽게 접할 수 있는 HMD 가 보급되며 일상생활에까지 영향을 미치고 있다. 이에 주로 쓰이는 영상은 좌우시점을 달리해 디스플레이하는 스테레오스코픽 방식이다. 이는 3D 비디오 영상을 만드는 방법 중 대표적인 방법으로, 좌우 시점간의 차이를 통해 입체적인 영상을 만들어내고, 몰입감을 더 높이는 것이다. 스테레오스코픽 영상을 만들기 위해서는 두 대의 RGB 카메라로 좌우시점을 분리해 촬영하는 방식과 RGB 카메라 한 대와 깊이 카메라 한대를 결합해 중간시점을 촬영하고 워핑을 통해 후에 좌우시점 영상을 만들어내는 방식이 있다. 깊이 카메라를 이용한 스테레오스코픽 방식은 두 대의 RGB 카메라 촬영보다 간편하며, 저장 데이터 용량도 더 적다. 이에 워핑을 이용한 스테레오스코픽 영상 제작 기법이 많이 연구되고있으며, 이를 위해 깊이 영상의 품질을 높이는 방법, 시점 이동 후 빈공간을 채우는 홀필링(Hole-filling)방법 등이 연구되고 있다[1-3]. 이에 맞춰 영상의 저작권을 보호하기 위해 워터마킹 방법이 필요하다. 하지만 워핑이라는 깊이 영상 기반 렌더링(DIBR, Depth Image Based Rendering) 기술을 사용하기 때문에 일반 2D 영상을 위해 연구된 워터마크와는 다른 워터마크 기술의 연구가 필요하다.

일반 2D 영상을 위한 디지털 워터마크 기술을 다양한 공격에 강인성을 갖는 것을 목표로 하고있지만, 현재 3D 영상을 위한 워터마크 방법은 깊이 영상 기반 렌더링 자체가 또 다른 형태의 공격으로 영상의 정보를 손상시키기 때문에 워터마크 추출이 더 큰 어려움을 겪고있다. 이에 깊이 영상 기반 렌더링을 위한 워터마크 연구가 많이 진행되어 있지 않다.

Wang 등이 제안한 방법은 SIFT(Scale Invariant Feature Transform)를 이용해 영역을 찾고 DCT 로 워터마크를 삽입하지만, 왼쪽시점 영상과 오른쪽시점 영상의 워터마크 추출 에러율의 차이가 크고, 평균 1.5%의 에러율을 보이고 있다[1]. 이 외에 H-D Kim 등이 제안한 방법은 DT-CWT 를 사용해 워터마킹을 수행하지만 5%이상의 에러율을 보이고 있고, Lin 등이 제안한 방법 또한 5%이상의 높은 에러율을 보인다[2,3]. 이처럼 3D 워터마킹 방법에 어려움이 많아 B-R Kim 등은 널블라인드(non-blind)방식으로 이동된 시점을 원본 시점으로 복원해서 워터마크를 추출하는 방식도 새롭게 제안한 적이 있다[4].

이에 본 논문에서는 2DDWT(2-Dimensional Discrete Wavelet Transform)를 이용해 워터마크(WM, watermark)을 중간시점 영상에 삽입하고 깊이 영상 기반 렌더링으로 좌시점 영상과 우시점 영상을 생성해서 WM 를 추출하는 블라인드(blind)워터마킹 방법을 제안한다. 본 방법은 DIBR 이라는 시점이동의 특수한 공격에 강인성을 가지기 위해 고안되었다. 이를 위해 1-Level DWT 를 수행하며, LH 부대역에 0 과 1 로 구성된 이진 워터마크를 암호화한 후 DWT 계수 양자화 기반으로 삽입한다.

2. 가상시점 영상 생성 방법

2.1 깊이 영상 기반 렌더링(DIBR)

깊이 영상 기반 렌더링 기술은 카메라의 기하학적 구조에 기반을 둔 방법이다. 이는 카메라의 내부, 외부 파라미터를 이용하여 기준시점의 RGB 영상의 좌표계를 깊이영상으로

실세계좌표(world coordination)으로 변환 했다가 가상시점의 카메라 파라미터를 이용하여 원하는 시점의 2 차원 공간으로 재투영하는 것이다.

카메라 파라미터와 영상 좌표값을 이용하여 3 차원좌표를 산출하는 과정은 식(1)과 같다.

$$\begin{aligned} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} &= K[R|T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = K^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} + T \\ &= R^T K^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} - R^T T = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1) \\ T &= \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \quad (2) \end{aligned}$$

여기서 x,y 는 영상 평면에 물체가 투영되는 2 차원 좌표이고, K 는 카메라 내부(intrinsic) 파라미터, R 과 T 는 카메라의 회전(rotation)과 이동(translation)에 대한 카메라 외부(extrinsic) 파라미터로 식(2)와 같이 표현한다. 그리고 X,Y,Z 는 투영되는 물체의 3 차원 좌표를 나타낸다. X 와 Y 의 좌표는 카메라 파라미터 계산식을 통해 구할 수 있지만 깊이 정보는 깊이영상을 통해 별도로 계산해야한다. 일반적으로 깊이 영상은 물체의 깊이를 0~255 의 값으로 변환하여 표현한다. 따라서 다음 식(3)를 통해 실제 3 차원 좌표의 Z 값을 산출한다.

$$Z(i,j) = \frac{1}{\left(\frac{D(i,j)}{255} \times \left(\frac{1}{Z_{MIN}} - \frac{1}{Z_{MAX}}\right) + \frac{1}{Z_{MAX}}\right)} \quad (3)$$

이를 통해 2 차원 공간으로 재투영하는 깊이 영상 기반 렌더링을 수행하면 그림 1 과 같이 나타난다. 이때 영상은 원본 시점에서는 보이지 않던 부분이 생기며 비폐색(dis-occlusion)영역이 발생하고 이를 그림 1 에 파란색 영역으로 표현하였다.

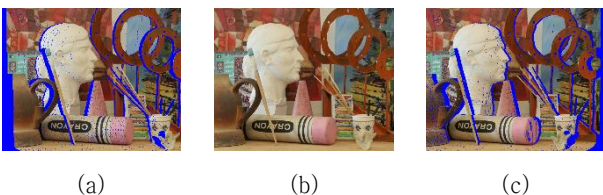


그림 1. 시점이동 예: (a) 왼쪽시점영상, (b) 중간시점영상, (c) 오른쪽시점영상

2.2 홀필링(Hole-filling)

홀필링은 렌더링으로 인해 생기는 비폐색(dis-occlusion)영역을 홀(hole)로 정의하고 그것들의 화소 값을

결정하는 방법이다. 홀은 시점이 변화하면서 객체에의해 가려져 있던 배경에 속하는 부분이다. 또한, 같은 객체내에서 깊이차이에의해 객체가 갈라지는 현상을 크랙(crack)이라 부르며 크랙도 홀필링으로 채울 수 있다. 이에대한 홀필링 방법은 많은 방법들이 연구되고 있지만, 본 실험에서는 수평적인 시점이동에대해 고려하기 때문에 수평적 탐색을 통해 배경에 속하며 홀이 아닌 픽셀을 찾아 홀을 채우는 방법을 사용한다. 배경인지 판단하기 위해 깊이 영상을 이용하며, 홀 근처에서 거리적으로 Z 값이 더 멀리 존재하는 픽셀을 판단하고 이를 이용해 홀필링을 수행한다.

그림 1 에서 보이는 비폐색영역을 홀필링한 영상을 그림 2 에 나타내고 있으며, 홀이 채워지며 자연스러운 영상이 된 것을 볼 수 있다. 하지만 좀 더 자연스러운 시점이동 영상이 되기 위해서는 더 정교한 홀필링 작업과 연구가 필요하며, 본 논문에서는 홀필링보다 워터마크 삽입과 추출에 더 집중하였기에 기본적인 홀필링에 멈추고 연구를 진행하였다.

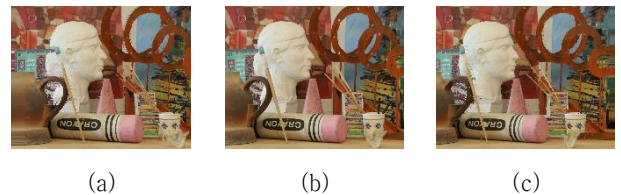


그림 2. 시점이동 후 홀필링 예: (a) 왼쪽시점영상, (b) 중간시점영상, (c) 오른쪽시점영상

3. 워터마크 삽입 및 추출

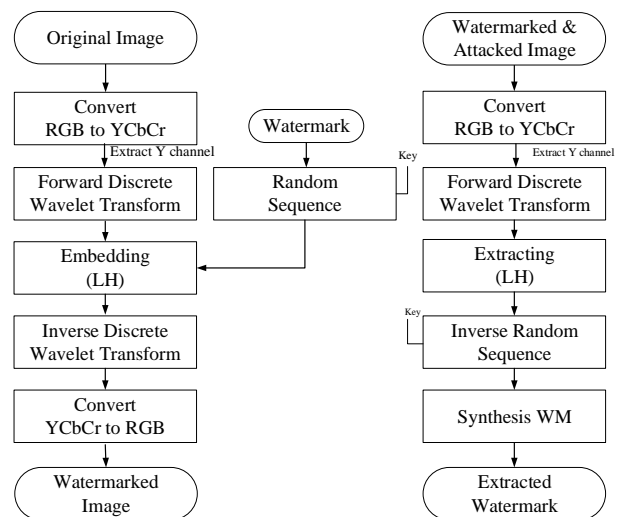


그림 3. 워터마크 삽입 및 추출 과정

WM 삽입 과정은 그림 3 에서 보는것과 같이 RGB 컬러영상을 강인성을 위해 YCbCr 형식으로 변환한다. 이후 Y 채널에 WM 를 삽입하며 이는 1-Level DWT 를 통해 주파수 도메인으로 변경된다. 주파수 도메인으로 변경하는 것 역시 강인성을 위한 방법이며, DWT 를 사용하는 것은 전역적이며 위치에 기반한 WM 삽입을 위한 것이다. 그림 4 에

1-Level DWT 후 변환된 영상의 모습을 나타내고 있으며, 좌측 상단부터 시계방향으로 각 부대역을 LL, HL, HH, LH 라 부른다. 이 중 LH 부대역을 사용한다. LH 의 의미는 가로방향으로 Low 하고, 세로방향으로 High 한 주파수 성분들의 집합을 의미한다. 이는 수평적으로 시점이 이동하는 랜더링의 경우 영상의 변화가 영향을 가장 적게 미치는 부대역이다. 따라서 수평적 시점이동 공격에 유리하다.



그림 4. Art 영상의 1-Level DWT 예

WM 삽입 위치를 선정하는 방법은 그림 5 와 같이 일정 간격을 두고 삽입하며, 1390x1110 크기 영상의 경우 총 19 개의 WM 가 삽입된다. 32x32 크기의 WM 를 16x64 크기로 변형하여 삽입한 것은 수평적 시점이동으로 변형되는 픽셀들 사이에서 데이터가 유지될 가능성이 큰 픽셀들이 세로적으로 분포하기 때문이다. 이는 현재 일시적인 실험으로 향후 연구에서 정확한 위치와 시점이동에 의한 변화 값의 상관관계를 분석한 후 개선할 예정이다.

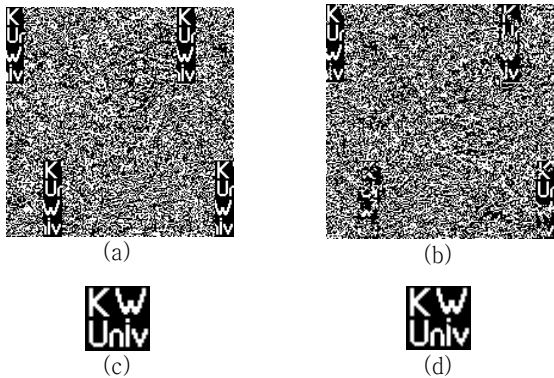


그림 5. (a) 중간시점 WM 삽입부분 일부 예시, (b) baseline distance 5% 오른쪽시점이동 영상 WM 추출부분 일부 예시, (c)삽입한 WM, (d)추출된 WM(BER=0.00293)

WM 데이터를 삽입하는 방법은 식(4)(5)와 같다. 여기서 W 는 DWT 영역에서 주파수 값이며, W' 은 워터마크 삽입 후 주파수 값, α 는 scaling factor, wm 은 삽입될 워터마크 0,1 을 의미한다. wm 은 k-stage LFSR(Linear Feedback Shift Register)를 통해 암호화된 WM 이다.

$$\begin{cases} \text{if } -\alpha < W < \alpha & \begin{cases} \text{if } wm = 0, & W' = +0.5\alpha \\ \text{if } wm = 1, & W' = -0.5\alpha \end{cases} \\ \text{else,} & W' = \alpha \cdot \left(\frac{W}{\alpha} + coef \right) \end{cases} \quad (4)$$

$$\begin{cases} \text{if } \left\lfloor \frac{W}{\alpha} \right\rfloor \bmod 2 = wm, & coef = 0.5 \cdot \frac{W}{|W|} \\ \text{if } \left\lfloor \frac{W}{\alpha} \right\rfloor \bmod 2 \neq wm, & coef = 1.5 \cdot \frac{W}{|W|} \end{cases} \quad (5)$$

식(4)에서 $\left\lfloor \frac{W}{\alpha} \right\rfloor$ 는 $\frac{W}{\alpha}$ 보다 작은 가장 큰 정수를 나타내고, $\bmod 2$ 는 2 로 나눈 나머지를 나타낸다. 식(5)는 $coef$ 를 판단하는 과정을 나타내었다.

WM 데이터를 다시 추출하는 방법은 식(6)과 같이 표현할 수 있다. wm' 은 추출된 워터마크이다.

$$wm' = \left\lfloor \frac{W'}{\alpha} \right\rfloor \bmod 2 \quad (6)$$

이렇게 나타난 추출영상에서 정확한 WM 를 찾는 방법은 기존에 삽입했던 WM 를 이용한다. 기존 WM 를 ReWM(Reference WM)이라하면, ReWM 을 추출영상에 투영시켜 추출영상에 있는 데이터와 ReWM 의 유사성을 측정한다. 측정방법에는 NCC(Normalized Cross Correlation)가 사용되며, 유사성이 완벽히 일치할 경우 그 데이터를 WM 로 판단한다. 하지만 유사성이 완벽히 일치하는 데이터가 추출영상에 존재하지 않는다면 일정 유사성 이상(NCC 0.5 이상)을 보이는 데이터를 추출하고, 이를 종합하며 최종 WM 를 판단한다. 데이터를 종합하는 방법은 0 과 1 의 이진 분포에서 더 많이 존재하는 값을 최종값으로 결정한다. 이렇게 결정된 최종값들이 모이면 최종 WM 를 판단할 수 있다. 이렇게 판단된 WM 가 그림 5 에 나타나 있으며, BER(Bit Error Rate)이 0.00293 으로 낮은 에러율을 보이고 있다.

4. 실험 및 결과

실험에 사용한 영상은 Middlebury Stereo Datasets[5]와 Microsoft Research 3D Video Datasets[6]에서 얻었으며, 컬러영상과 깊이영상을 종류별 1 개씩 사용하였다. 사용한 영상 종류는 표 1 에 나타내었다. 사용된 영상의 크기는 1390x1110, 1342x1110, 1366x1110, 1372x1110, 1024x768 이다. Middlebury Stereo Datasets 에서 얻은 영상은 view1 깊이영상과 Exposure1 의 컬러영상을 사용했으며, Microsoft Research 3D Video Datasets 에서 얻은 영상은 동영상의 첫번째 프레임에 해당하는 컬러영상과 깊이영상을 사용하였다. WM 는 원본 컬러영상에 삽입되고, 이를 좌시점, 우시점으로 랜더링 한 후 이동된 시점에서 WM 을 추출하는 실험을 진행하였다. 추출되는 WM 는 BER 로 에러 정도를 나타내었다.

실험은 총 10 개의 영상에서 얻은 값을 평균해서 기록하였다. 영상에 삽입되는 WM 는 0 과 1 로 구성된 32x32 크기이다. 이는 1024bit 를 삽입하는 것과 같다.

표 1. 실험에 사용된 10 개의 컬러영상과 깊이영상

영상 크기	개수	종류
1390x1110	4	Art, Dolls, Moebius, Books
1342x1110	2	Reindeer, Laundry
1366x1110	1	Midd
1372x1110	1	Wood
1024x768	2	Ballet, Breakdance

표 2. WM 삽입 후 비가시성 비교(평균 PSNR, 평균 SSIM)

	PSNR [dB]	SSIM
Proposed method	61.74	0.9999
Wang's method[1]	59.53	0.9990
Kim's method[2]	42.15	0.9890
Lin's method[3]	42.55	0.9930
Lin's method[3]	42.43	0.9900

그림 6 을 보면 WM 삽입 후에도 영상의 비가시성이 좋다는 것을 알 수 있다. 이는 표 2 에서 나타나 있듯이 평균 PSNR 61.74dB, SSIM 0.9999 로 수치적으로도 뛰어난 비가시성을 보여주고 있다. 일반적으로 PSNR 40dB 이상의 영상은 시각적으로 구분할 수 없을 정도로 비슷하다고 알려져 있으므로 이는 WM 삽입된 영상이 원본 영상과 유사성이 크다는 것이며, 다른 기존논문들 또한 PSNR 40dB 이상의 유사성을 보여주고 있다. 그리고 WM 추출율을 비교하면, 본 논문에서 제안하는 방법이 다른 기존논문들보다 현저히 낮은 에러율을 보이는 것을 표 3 에서 확인할 수 있다. 또한, Wang 등이 제안한 방법은 왼쪽시점 영상과 오른쪽시점 영상의 에러율 차이가 2.07%로 본 논문에서 사용한 방법이 0.08%차이를 보이는 것에 비해 편차가 크다. 하지만 본 논문이 고안한 WM 삽입 및 추출 방법은 깊이 영상 기반 렌더링을 고려한 방법으로 2D WM 삽입 및 추출 방법처럼 발전하기 위해서는 필터링과 압축 및 기하학적 공격에도 강인성을 가지도록 다양한 실험과 추가적인 개선이 필요하다.



그림 6. 비가시성 예: (a) 원본영상, (b) 워터마크된 영상

표 3 가상 시점 WM 에러율 비교(평균 BER)

	Center image	Left image	Right image
Proposed method	0.000	0.000879	0.001758
Wang's method	0.000	0.005100	0.025800
Kim's method	0.009	0.023000	0.034000
Lin's method	0.000	0.068000	0.062000
Lin's method	0.006	0.111000	0.094000

5. 결론

본 논문에서는 고부가가치 디지털 콘텐츠인 양안식 또는 자유시점 영상을 위한 3D 영상 워터마크 삽입 및 추출 알고리즘을 제안했다. DWT 를 이용한 방법으로 기존에 LL, HL, LH, HH 부대역에 WM 를 삽입하던 것과 달리 시점이동 영상 특성에 맞게 WM 를 삽입하는 방법의 일부를 실험하였고, 좋은 결과를 보였다. 수평적 시점이동에는 LH 성분 에 WM 를 삽입하는 것이 좋은 효율을 보이는 것을 확인했으며, 적은 WM 데이터를 삽입하더라도 효과적으로 낮은 에러율을 보이며 추출가능하다는 것 또한 확인하였다. 이로써 향후 연구를 통해 3D 디지털 콘텐츠 저작권을 보호할 워터마크로 충분히 활용가능 할 워터마크 삽입 알고리즘이라 기대된다.

감사의 글

이 연구는 2014 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2013R1A1A2057798)

참고문헌

[1] Shen Wang, and Chen Cui, "Watermarking for DIBR 3D images based on SIFT feature points." Measurement, Vol. 48, pp.54-62, Feb. 2014.

[2] H-D Kim, and J-W Lee, "Robust DT-CWT Watermarking for DIBR 3D Images." IEEE Transactions on Broadcasting, Vol. 58, No. 4, pp.533-543, Oct. 2012.

[3] Y-H Lin, and J-L Wu, "A Digital Blind Watermarking for Depth-Image-Based Rendering 3D Images." IEEE Transactions on Broadcast, Vol. 57, No. 2, pp.602-611, Jun. 2011.

[4] B-R Kim, and D-W Kim, "Digital Watermarking on Image for View-point Change and Malicious Attacks." Journal of Broadcast Engineering, Vol. 19, No. 3, pp.342-354, May. 2014.

[5] <http://vision.middlebury.edu/stereo/>

[6] <http://research.microsoft.com/en-us/downloads/5e4675af-03f4-4b16-b3bc-a85c5bafb21d/>