

3차원 스테레오 영상 부호화 기법

□ 유지상 · 김은수 · 배진우 · 이병진 / 경운대학교 전자공학과

1. 서론

최근, 동영상 압축에 대한 연구는 MPEG-1/2/4 또는 H261, H263 등의 압축 표준의 제정으로 하나의 결실을 맺게 되었고[12], 현재도 지속적으로 연구, 개발되고 있다. 이에 더하여, 인간의 끊임없는 욕구는 보다 사실적이고, 현장감 있는 영상을 추구하게 되었고, 이에 따라 3차원 영상의 정보처리에 관한 연구가 새로운 주제로 등장하여 다양한 분야에서 활발히 연구되고 있다. 이와 관련하여 선진국에서는 HDTV 개발과 더불어 영상의 현실감과 입체감을 주기 위한 3차원 입체 TV에 관한 연구 개발에 노력을 기울이고 있다[1]. 3차원 입체영상의 신호처리기술, 압축기술, 전송기술 및 디스플레이 기술이 개발되면, 3차원 영상전송 및 3차원 입체 TV 방송은 HDTV를 보완 또는 대체하는 차세대 기술로 발전해 나갈 뿐만 아니라, 3차원 부

조지도 제작에 대한 항공사진, 의료영상, 원격조정, 현미경학과 같은 여러 분야에 응용이 가능할 것이다. 또한, 일본 NTT에서 주장하는 VI&P(visual intelligent and personal) 즉, 가시성, 지능성 및 개인성이 추가되어 인간에게 실제감과 친밀감을 주는 정보통신 서비스가 구현될 것으로 전망되고 있다[3]. 이러한 가시성의 추가란, 인간이 주변의 세계로부터 받아들이는 정보의 80% 이상이 눈을 통해 이루어진다는 사실에서 필연적으로 구현되어야 하며, 인간이 현실에서 느끼는 것과 같은 현실성과 자연성이 뛰어난 영상통신의 구현을 위해서는 입장감을 극대화 할 수 있어야 한다. 입장감의 극대화는 결국 인간 생활공간의 3차원화에 기인한 영상의 3차원적 디스플레이를 의미한다. 따라서, 이러한 초고속 정보통신망의 기반 위에서 실현될 서비스들은 현재의 “보고 듣고 즐기는” 멀티미디어형 서비스로부터 “실감 있게 보고

듣고 느끼는” 3차원적 입체 멀티미디어 서비스가 구현될 것으로 전망되고 있다[4].

이러한 3차원 입체영상은 인간의 시각시스템(human visual system : HVS)에서 기인한다. 인간이 물체를 입체로 인식할 수 있는 것은 오른쪽 눈과 왼쪽 눈이 각각 다른 영상을 인식하기 때문이다. 이 두개의 다른 영상을 뇌가 고속처리 함으로써 인간은 입체감과 원근감을 느낄 수 있다[56]. 이러한 원리를 이용하여, 양쪽 눈에 스테레오 카메라로 촬영한 각각 다른 영상을 보여주어 인위적으로 인간에게 원근감을 제공할 수 있다. 이것이 스테레오 영상인데, 이런 방식으로는 입체시점이 단 하나로 제한되며, 카메라 설정과 관찰자의 인체 구조의 차이(눈 사이의 거리차) 등에 따라 인식되는 거리감에도 차이가 발생하며, 경우에 따라서는 거리감 인식이 불가능하거나 심한 피로감을 느끼게 된다. 이에 대한 보완책으로, 다시점 카메라로 획득한 영상(다시점 영상)을 이용함으로써 시점을 확장하는 동시에 관찰자에게 적절한 영상을 제공할 수 있게 된다[78]. 그러나, 스테레오 영상 전송과 관련해서 가장 중요한 문제점은 두 화면이 각각 독립적으로 부호화 된다면, 단안 화상을 전송하는 것과 비교해 볼 때, 스테레오 화상전송에 필요한 대역폭은 두 배로 증가하게 된다는 것이다. 또한, 향상된 성능과 “Look-Around” 효과를 제공하기 위한 다시점 영상의 경우에는 결국 몇 배의 전송 대역폭을 필요로 하게 된다. 이러한 이유에서 영상정보의 압축기술은 3차원 영상처리에 있어서 가장 중요한 요소중의 하나로 인식되고 있다.

따라서, 본 논문에서는 전술한 3차원 영상시스템의 핵심기술이 되는 효율적인 3차원 스테레오 영상 부호화 기법에 관하여 살펴보고자

하며, 그 구성은 다음과 같다. 2절에서는 스테레오 영상을 어떻게 인간의 시각 시스템이 인식 할 수 있는가에 대한 고찰로써 이론적인 배경과 스테레오 영상 부호화에 대한 전반적인 이론을 살펴본다. 3절에서는 스테레오 영상 전송시스템에서 핵심이 되는 변이 보상 예측방법에 대해 살펴 보고, 이때 사용되는 변이 추정에 관하여 기존에 제안된 알고리즘을 중심으로 설명한 후, 차 영상에 대한 효율적인 부호화 방법에 대해 설명하겠다. 마지막으로 4절에서는 결론 및 향후 연구방향을 소개한다.

2. 이론적 배경

본 절에서는 스테레오 영상처리에 대한 기본적인 이론을 소개한다. 2.1절에서는 인간이 시각을 통해 어떻게 깊이감을 인식하는지에 대한 인간 시각계의 기하학적 분석에 대해 알아보고, 이를 통하여 두 대의 카메라로 얻어진 스테레오 영상의 기하학적 의미를 살펴본다.

2.1 기하학적 고찰

우리의 두 눈이 자연계의 정경에서 어떻게 깊이감을 인식하는지에 대한 기하학적 고찰을 통하여 3차원 디스플레이 시스템에 대해 보다 잘 이해 할 수 있다.

본 절에서는 이러한 인간 시각 시스템(human visual system : HVS)의 구조를 살펴보고, 이러한 인간의 시각 시스템과 카메라의 구조를 기하학적으로 해석한 에피폴라(epipolar) 기하학에 대하여 살펴보겠다.

2.1.1 인간 시각계의 기하학적 분석

인간의 두 눈은 가로방향으로 65mm 정도

이격 되어있다. 이러한 두 눈으로 물체를 바라봄으로써 망막에 맺힌 한 쌍의 2차원 신호를 통하여 물체의 거리와 공간을 인식하고 외형적인 입체 구조를 감지한다. 서로 다른 점에 대한 눈의 반응을 표현한 그림 1을 살펴보면, 왼쪽과 오른쪽 두 눈이 공간상에서 각 θ 를 형성하면서 한 점 P1을 응시하는 것을 보여주고 있다. 또한 점 P2에서는 P1과 공간적 깊이차를 가지는 영상을 만들어 낸다. 그리고 이들 영상은 각각의 망막에서 다른 거리에 위치하게 된다. 즉, 서로 다른 값을 가지게 되는데 이 각의 합을 망막시차(retinal disparity)라고 한다. 묘사한 바와 같이 눈의 안쪽을 향하도록 측정된 망막의 각은 양수값을 갖는다. 두개의 이미지가 함께 합쳐질 때 두뇌에 물체에 대한 깊이와 형태에 대한 정보를 제공하는 역할을 하는 것이 이 망막 시차이다. 또한, P1을 주시할 때 생기는 각 α 를 형성하게 되는데, 이를 폭주각(vergence angle)이라 한다. 여기서 명확히 알 수 있는 것은,

$$\alpha + \delta_1 + c + \delta + d = 180$$

이고, 따라서

$$\beta + c + d = \delta_1 - \delta_2$$

이 성립한다. 식(2.1)에서 볼 수 있듯이 양안시차는 $\alpha - \beta$ 라는 식으로 표현 할 수 있다. 이는 실제 광경내의 두 점에 의해서 생긴 폭주각 사이의 차는 망막 시차의 크기와 같다는 것을 의미한다. 일반적으로 $\beta - \alpha$ 일때 망막시차는 음수가 된다. 점 P2에 초점을 맞추기 위해서는 눈들이 교차해야 하기 때문에, 이것을 교차 시차(crossed disparity)라고 한다. 이에 반해, 망막시차가 양수이면 비교차 시차(uncrossed

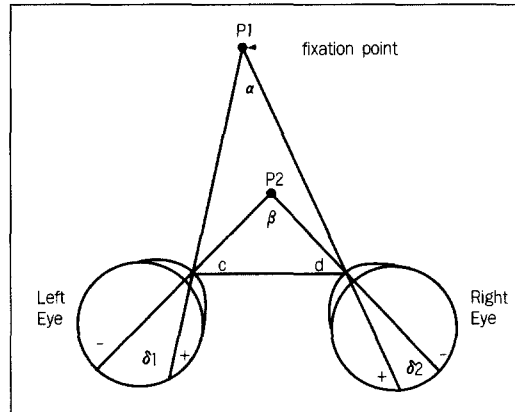


그림 1. 시각계의 기하학적 설명

disparity)라고 한다[9]. 여기서 내리는 중요한 결론은 우리의 두뇌가 인식하는 깊이감은 망막시차와 관계가 있다는 것이다. 예컨대, P1과 P2의 깊이감이 커지면 폭주각의 차도 커지게 되며, 결국 망막시차의 크기도 커지게 됨으로써 우리의 뇌는 깊이감을 인지 할 수 있게 된다.

일반적인 스테레오영상 디스플레이에서는 이 효과가 가장 많이 활용되고 있으며, 대뇌에서도 이 같은 양안 시차를 검출하는 세포가 존재한다는 것이 알려져 있다.

2.1.2 에피폴라(Epipolar) 기하학

스테레오 영상 비디오 시스템은 동일한 실제 장면을 관찰하는 2개의 카메라로 구성되어있기 때문에, 촬영되는 영상에 몇 가지 제약적인 요소를 내포하고 있다. 즉, 주어진 한 쌍의 2차원 좌표로부터 3차원 좌표를 만들기 위해서는 각 영상간의 블록 중 대응에 관한 일치 문제(correspond problem)를 해결해야 한다. 이러한 기본적인 제약사항은 그림 2와 같은 에피폴라 기하학(epipolar geometry)[10]으로 해석되고 있다. 에피폴라 평면(epipolar plane)은 공

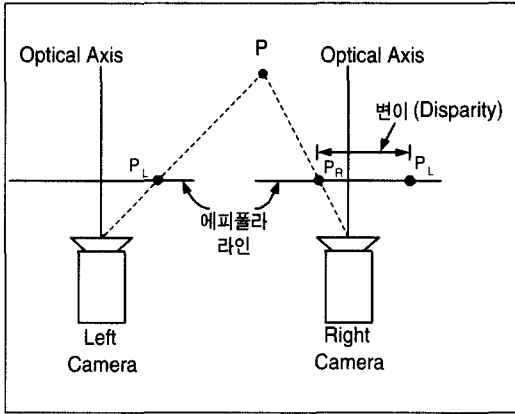


그림 2. 스테레오 카메라 에피폴라 기하학의 접목

간상에서 점 P와 초점거리 F를 가진 두 렌즈의 중심을 연결하는 선에 의해 정의된다. 또한 이 평면과 각 투영 영상이 이루는 교선을 에피폴라 선(epipolar line)이라 한다. 에피폴라 기하학이란, 특정한 에피폴라 평면의 모든 점들은 영상 평면의 에피폴라 선상의 해당하는 점에 대응된다는 기하학적 원리이다. 이 때, 점 P에 대한 P_L 과 P_R 의 위치차를 변이(disparity)라 하며, 이를 벡터로 표시한 것이 변이 벡터이다.

이러한 원리를 이용하여, 에피폴라 기하학은 모든 스테레오 영상 매칭 방법의 기본이 된다.

에피폴라 기하학을 기반으로 하여 인간의 시각과 같은 구조로 카메라를 제작하여 촬영하면 스테레오 영상을 인위적으로 만들 수 있다. 이러한 카메라를 스테레오 카메라라고 하며 스테레오 카메라는 기하학적 구조에 따라 평행식, 평행이동식, 폭주식과 같이 세가지 구조로 나누어 볼 수 있다[11]. 평행식 구조의 경우, 에피폴라 선이 두 영상에 평행하므로, 변이벡터가 항상 수평 값만을 가진다. 하지만, 폭주식 구조의 경우 좌, 우 영상의 에피폴라 선이 서로 평행하지 않기 때문에 변이벡터는 수직, 수평성분을

모두 가진다. 따라서, 수평변이를 갖는 스테레오 영상으로 분석하기 위해서는 이를 교정하는 전처리 과정이 요구된다[12,13].

3. 스테레오 영상 압축 알고리즘

스테레오 영상 압축 방법 중 가장 많이 사용되고 있는 방법이 변이 추정을 이용한 방법이다. 이 방법은 스테레오 영상 쌍 중 기준이 되는 영상을 독립적으로 부호화하고, 다른 영상은 기준이 되는 영상에서의 변이를 추출 후 이를 보상하는 방식이다. 추정 후 추정된 영상과 실제 영상의 차이를 구하게 되고, 여기서 구해진 차이 영상이 오차영상(residual image)이 된다.

이렇게 함으로써 스테레오 영상의 높은 공간적 중복성을 제거하고, 변이추정을 거쳐 나오는 변이 벡터를 수신단에 전송하여 추정된 변이 벡터를 통해 예측된 값과 실제 영상의 차이를 변환 부호화해서 높은 압축 효율을 얻을 수 있다. 그림 3은 변이 보상 방법을 블록 다이어그램으로 표시한 것이다. 3절에서는 스테레오 영상 전송시스템과 여기에 사용되는 변이추정방

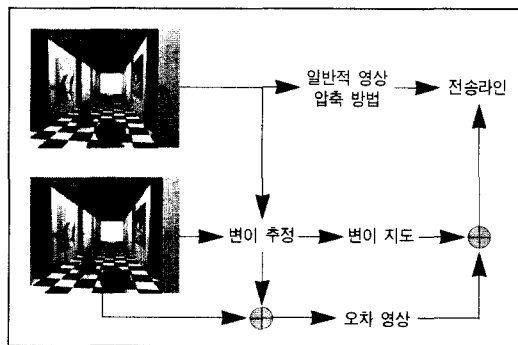


그림 3. 스테레오 영상의 변이 보상 기반 부호화

법을 제안된 논문 위주로 살펴 본 후, 변이 추정 후 얻어지는 차 영상 부호화의 방법에 대해 설명하겠다.

3.1 변이 추정(Disparity Estimation)

변이 추정이란, 스테레오 영상의 좌, 우 영상이 공간적 상관성이 매우 높다는 점을 이용하여 서로 대응되는 점 또는 영역을 찾아내는 과정이다. 이렇게 함으로써 두 영상간의 공간적 중복성을 제거하고 압축효율을 높이는 방식이다. 스테레오 영상 압축에 관한 논문은 단안 영상이나 동화상 압축에 비해 매우 드문 편이다. 그리고 그 중 대다수의 논문들이 변이 추정의 성능을 높이는 방법에 초점을 맞추고 있다. 변이추정은 그 추정 방향에 따라 순방향 추정(forward estimation)과 역방향 추정(backward estimation)으로 나누어 볼 수 있다[14,15]. 스테레오 영상 중 좌영상을 기준영상으로 정했을 때, 순방향 추정은, 좌영상을 분할 또는 특징점을 추출 한 후, 우영상에서 가장 유사한 영역을 찾아 좌영상으로 우영상을 추정해 내는 방식이다. 반대로, 역방향 추정은 생성할 영상인 우영상을 분할하거나 특징점 추출 후 기준 영상에서 이의 정합점을 찾아나가는 방식이다. 정방향 추정의 경우 좌영상의 분할 영역정보 혹은 특징점에 관한 정보 없이, 변이 정보만 전송하면 수신단에서 수신된 좌영상을 이용하여 송신단에서와 같은 알고리즘으로 분할 영역 정보나 특징점 정보를 동일하게 찾아낼 수 있어 압축효율을 높일 수 있는 장점이 있지만 송신단에서 계산량이 증가하고, 기준영상이 올바르게 복원되지 못했을 경우 오차가 영상전체에 퍼져 나타나게 되는 단점이 있다. 반면에 역방향 추정의 경우에는 수신단의 복호화 과정이 간단

한 대신 분할 영역이나 특징점들에 관한 정보를 함께 전송해야 하므로 압축율이 낮아지게 된다[16]. 여기서는 일반적으로 사용되는 시차 추정 방법 중에서 블록 기반, 객체 기반, 망 기반 방식에 대하여 알아보도록 하겠다.

3.1.1 블록기반 변이추정(Block-based Disparity Estimation)

블록 기반 방식은 영상을 고정크기, 혹은 가변 크기의 블록으로 나누어 변이를 추정하는 방식이다. 이는 MPEG-1/2, H.261, H.263 등의 영상 압축 표준에서 16(16의 매크로 블록(macro block: MB)단위로 움직임 추출 시 사용되는 기법과 동일하다[17]. 스테레오 영상 중 한 영상을 블록 단위로 나누어 다른 한 영상에서 비용 함수(cost function)를 만족하는 위치를 찾아 이 이동량을 그 블록의 변이 값으로 결정한다. 비용함수는 일반적으로 MAE, MSE, NTD, 상관도(cross correlation) 등이 있다. 하지만 MSE의 경우에는 하드웨어적으로 제곱연산 구현의 복잡성 때문에 잘 활용되지 않고있다. 아래 식은 블록정합 방식에서 비용함수를 MAE를 이용했을 때의 식이다.

$$MAE(i, j) = \frac{1}{N_x N_y} \sum_{k=1}^{N_x} \sum_{l=1}^{N_y} |I_l(k, l) - I_l(k+i, l+j)|$$

여기서, I_l 은 예측하고자 하는 영상이고, I_1 은 참조영상을 말한다. 먼저, 예측 추정하고자 하는 영상을 $N(M$ 크기의 블록으로 분할 후, 검색구간 내에서 최소의 MAE(mean absolute error)를 얻는 블록이 가장 상관도가 높은 블록으로 판별된다. 그림 4는 두장의 스테레오 영상에서 좌영상을 기준 영상으로 하고 우영상의 탐색 구간 내에서 최대 매칭 블록을 찾아내는

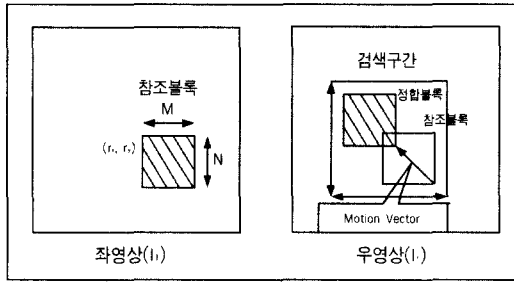


그림 4. 블록기반 변이 추정 방식

과정이다. 위의 설명에서는 기준 블록의 변이가 수직, 수평 성분을 모두 가지는 경우이고, 일반적인 스테레오 영상에서는 스테레오 카메라가 병렬로 놓여있다고 가정하기 때문에 수평시차만 고려하여 계산 시간을 단축하는 경우도 있다[16,18,19]. 블록기반 변이 추정방법은 구현이 간단하다는 장점이 있지만 블록의 크기와 탐색 범위에 따라 변이추정 결과에 많은 영향을 미친다. 즉, 블록의 크기가 커지면 다른 깊이로 가지는 물체가 동시에 블록 내에 존재할 수 있으므로 변이 벡터가 부정확해지고 블록화 현상이 일어날 수 있으며, 블록의 크기가 작으면 세밀한 변이 벡터를 얻을 수 있는 반면, 일치점이 중복되어 신뢰도가 감소한다. 탐색범위가 넓어지면 여러 후보 블록 중에서 대응 블록을 찾을 수 있지만 계산량이 증가하여 실시간 처리에 힘든 단점이 있다. 반면에 탐색범위가 적어지면 변이 벡터의 정보량이 감소하고, 검색시간을 줄일 수 있지만, 예측된 영상의 화질열화 현상이 발생한다.

이러한 블록의 크기와 탐색범위로 인한 잘못된 변이 추정을 보완하기 위해 Sathurama는 가변블록을 기반으로 하는 변이 추정방법을 제시하였다[18]. 그는 일정한 밝기 값을 가지는 영역끼리 분할하여 각 영역의 변이를 추정하는

방법을 사용하여 경계와 가려진 영역에서도 비교적 정확한 변이추정이 가능하게 하였지만, 하지만 분할된 블록의 크기가 각각 다르기 때문에 이에 대한 추가적인 정보를 전송하여야 하는 단점이 있다. 또한 Tzovaras와 Strintzis는 계층적 블록 정합방식(hierarchical BMA : HBMA)을 이용하여 탐색범위를 한정하여 수행시간을 단축시키는 방법을 제시하였다[19]. 이 방법에서는 좌, 우 영상을 웨이블릿 패킷(wavelet packet) 분해하여 각각 저역 통과 필터와 다운 샘플링 과정을 통한 후 각각 다른 해상도의 영상을 얻는다. 최저해상도의 영상에서 변이 추정을 시작하여 초기 변이 벡터를 구하고, 이때 계산된 변이 벡터는 한 단계 고해상도 영상의 변이 벡터의 초기값이 되어 이를 기준으로 다시 변이 추정을 하여 점차 높은 해상도의 영상으로 옮겨가면서 최종적인 변이추정을 하게 된다. 이때 한 단계 고해상도로 옮겨지면서 탐색범위를 한정하여 수행시간을 단축시킬 수 있다. Kim과 Ahn은 FS(Full Search) 방법의 변형으로 검색 시 방향성과 검색구간에 제한을 두어 계산 효율을 증가시키는 ADLS(adaptive directional, limited search)라는 방법을 제시하였다[20]. 이들은 스테레오 영상의 특징 중, 왼쪽영상을 기준 영상으로 하고 순방향 변이 추정을 할 때 대다수의 변이벡터가 왼쪽으로 이동 한다는 점에 착안하여 검색 방향의 우선순위와 이웃블록과의 변이 연속성을 이용하여 변이 추정의 효율적인 계산을 제안 하였다.

3.1.2 특징 기반 변이추정

블록기반 변이 추정방식은 두 영상에서 화소들의 밝기값으로 대응블록의 변이를 찾는다.

화소들의 밝기값만 가지고 대응블록을 찾으면 카메라의 광도변화나 물체의 경계 부분에서 블록기반 변이 추정방식은 효율이 떨어진다. 특징기반 변이 추정방식은 밝기값 대신 영상의 특징을 사용하여 대응블록을 찾는 방식이다. 이때 사용되는 영상의 특징으로는 영점 교차(zero-crossing)의 부호[21], 경도의 첨두(gradient peak)[22], 영역(region), 선(line), 경계(edge)등이 있다. 위의 특징들은 정합을 위해 독립적으로 사용되기도 하지만 블록 정합 방식과 같이 물체 윤곽에 따른 변이의 연속성이나 이웃 블록의 변이 연속성에서 제약이 가해지는 정합 방식의 단점을 보완하기 위해 같이 쓰는 경우도 있다. Kim과 Shon은 블록기반 변이추정 방식에서 변이 지도의 연속성과 정보량의 감소를 위해 주변블록과 변이 벡터를 평활화하는 방법을 사용하였는데 이때, 경계영역에서는 그 방향으로는 변이를 평활화하지 않는 경계보존 방향성 평활화 방식을 제안하여 블록기반 변이 추정 방식과 특징 기반 변이 추정 방식을 접목시킨 알고리즘을 제시하였다[23].

3.1.3 그 밖의 변이추정 방식

그 밖의 시차 추정 방식으로는 메쉬(mesh) 기반, 객체 기반 시차 추정방식 등이 있다. 이들 방식들은 블록기반 시차 추정 방식의 단점을 보완하고자 제안된 방법 들이다. 메쉬 기반 변이 추정 방식은 변이차가 있는 물체를 블록이 아닌 입의의 형태로 분할하여 각 분할 영역에 대해 평행 이동만으로 변이를 보상하는 대신 회전, 확대 축소, 변형 등의 변이를 처리하는 공간변환[24]을 채택하여 변이를 보상한다. 즉, 영상을 다각형 조각(polygonal patch)으로

나누어 다각형 조각의 꼭지점들을 중심으로 이동점을 찾아 변이를 추정하는 방식을 메쉬 기반 변이 추정 방식이라 한다[24,25,26]. 메쉬 기반방식은 메쉬의 형태를 결정하는 특정한 과정이 필요하며, 절점 정보만으로 불규칙 삼각망을 생성하는 알고리즘인 델로니 삼각망(delaunay triangulation)형성 방식이 대표적이다[25].

객체기반 방식은 영상이 독립적인 객체의 조합으로 이루어져 있다는 가정하에 영상을 객체 단위로 분할하여 변이를 추정하는 방식이다. 즉, 영상을 유사한 특성을 갖는 영역으로 분할(segmentation)하여, 분할된 영역에 대하여 변이를 찾고, 이와 함께 윤곽선(shape)을 부호화 하여 전송한다[15]. 객체단위로 변이를 추정할 경우, 블록화 현상이 발생하지 않으며 영상을 객체단위로 접근하는 것이 가능하므로 MPEG-4에 적합하고, 객체 형태 등의 추가 정보를 전송해야 하며, 객체가 세분화될 경우 객체 형태를 위한 데이터량이 증가하는 단점이 있다. 객체 기반 방식에는 영상 분할을 통해 VOP(video object plane)를 추출한 후, 각 VOP에 대한 윤곽선과 텍스처를 부호화하여 전송하는 방식이 적용된다[26,27].

3.2 차 영상 부호화

좌, 우 두 개의 영상이 동시에 전송되어야 하는 스테레오 영상 통신에서는 먼저 좌, 우 영상의 공간적 상관성을 이용한 변이 추정 및 보상을 하게 된다. 추정된 변이와 함께 예측된 영상과 원 영상과의 차 영상을 전송하여 효율적인 전송을 하게 된다. 특히 차 영상은 가려진 영역(occlusion region)의 정보를 가지고 있기 때문에 그 정보는 인간의 눈이 입체로 느끼는데 꼭 필요한 정보를 포함하고 있다. 따라서, 입체 영

상 시스템에서 차 영상이 차지하는 비중 또한 크다고 할 수 있다. 이번 절에서는 일반적인 DCT를 기반으로 하는 JPEG 표준 알고리즘을 살펴보고, 좀 더 효율적인 알고리즘을 제안된 방식 위주로 살펴보겠다.

3.2.1 DCT기반의 JPEG표준

JPEG(Joint Photographic Experts Group)은 DCT를 기반으로 하여 컬러 및 흑백 정지영상의 압축방법에 대한 표준안이다. JPEG은 압축율에 따른 화질의 상태를 고려하여 사용자의 목적에 따라 화질과 압축율을 선택할 수 있도록 하였다.

DCT는 고속 알고리즘을 갖는 직교 변환으로 에너지 집중 현상을 이용하여 보다 효과적인 압축을 할 수 있다. 특히 이차원 DCT의 기저함수는 분리성이 성립하고 대칭적이므로 두 번의 일차원 DCT를 수행하여 구할 수 있다. DCT의 고속처리 알고리즘도 하드웨어로 구현하기 위해 여러 알고리즘 구조에 대한 연구들도 진행되고 있다. JPEG 알고리즘은 정보량을 줄이기 위해 시각적으로 무시할 만한 정도의 화질 손상을 감수하고 AC계수들에 대해 JPEG에서 제공하는 양자화 테이블을 사용하여 양자화 과정을 수행하는데 양자화는 DCT계수를 양자화 테이블의 크기로 나누었을 때 가장 근접한 정수로 정의한다. DCT기반 부호화 과정의 마지막 단계는 엔트로피(entropy) 부호화 과정이다. 엔트로피 부호화는 DCT 계수들의 통계적인 특성을 바탕으로 좀더 압축할 수 있도록 한다.

DCT계수 중에서 DC 계수의 부호화는 한 블록의 DC 계수와 그와 인접한 블록의 DC 계수의 차를 허프만(Huffman) 부호화 테이블의

해당 코드로 할당한다. AC 계수는 부호화를 위해 지그재그(zigzag)방식으로 스캔한다. 스캔된 AC계수를 RL(run length)부호화하여 이에 따라 허프만 코드를 할당한다. 이에 대해 좀 더 효율적인 차 영상의 부호화를 위해 영상의 경계특성에 따라 AC 계수들의 분포를 이용한 방법도 있다[28]. 이 방법은 수평적인 경계가 많은 입력에서 AC 계수의 분포가 열의 방향으로 나타나며, 수직적인 경계가 많은 입력에는 행의 방향으로 계수가 분포하게 되고 대각 경계가 많은 입력에는 대각선의 방향으로 계수가 분포한다는 성질을 이용하여 부호화 방법을 개선하였다.

3.2.2 웨이블릿을 이용한 차 영상 부호화

기존의 DCT 기반의 영상 압축은 압축율을 높이면 블록화 현상으로 인한 영상의 열화현상이 나타난다. 이러한 단점을 보완하기 위해 웨이블릿(wavelet)을 통해 압축을 하는 방법이 있다[4]. 웨이블릿을 이용한 영상압축은 높은 압축율에서도 영상의 블록화 현상이 없어 주관적 화질이 좋고, 점진적(progressive) 영상 전송이 가능하다. 웨이블릿 변환을 통해 계층적인 영상구조로 원영상을 부호화 하는 경우 영상 자체의 고유 성질인 자기 유사성(self-similarity)을 이용하여 부호화 효율을 높일 수 있다[29]. 웨이블릿 변환 후 사용되는 대표적인 양자화 방법에는 SPIHT(Set Partitioning in Hierarchical Trees)[30]가 있는데, 이는 다양한 제로 트리 기법으로 영상을 부호화 할 수 있다. SPIHT 알고리즘의 기본 개념은 웨이블릿 계수들을 크기와 비트 평면(bit plane)에 의해 정렬하고, 집합 분리 배열 알고리즘에 의한 순서대로 전송한다. 계

$$n = \lceil \log_2(\max_{(i,j)} \{|C_{i,j}|\}) \rceil$$

수들의 부호화는 위 식에서 n 을 임계값으로 선택하여 n 을 감소시키면서 분류 과정과 세분화 과정에 의해 이루어지게 된다.

목록에 있는 계수들의 주요 계수 양자화 방법은 연속 근사화 스칼라 양자화(successive approximation scalar quantization)를 사용해 양자화 하는 방법이다. 이러한 웨이블릿을 이용한 효율적인 영상 압축을 인정하여, JPEG2000과 MPEG-4에서도 부분적으로 채택하고 있다[31].

4. 결 론

본 논문에서는 초고속 정보통신망이 구축되면서 3차원 영상 산업분야 즉, 화상 통신, 입체 영상회의, 입체 방송 등과 같은 3차원 데이터의 시각화에 관심이 고조되고 있는 시점에 발맞추어 3차원 영상통신에 필수적인 요소라고 할 수 있는 3차원 영상 압축 방식에 대하여 살펴 보았다. 스테레오 영상을 압축하기 위하여 시차 추출, 시차지도 압축, 차 영상 부호화로 나누어지는 인코딩 부분과 역순의 과정을 거치는 디코딩 부분을 전체적으로 다루었지만, 자세한 내용은 참고문헌으로 대신하였다. 결론에서는 본 사례 연구에 대해 간략하게 요약하겠다.

첫번째로 휴먼팩터(human factor), 즉 인간의 시각특성과의 정합을 취하는 것이 가장 중요한 과제이다. 입체 영상을 보았을 때에 생기는 특유한 왜곡, 예컨대 perfect theater 효과, 카드보드(card board) 효과[3]의 시각기구의 해명과 경감책, 안정피로의 원인해명과

해소책 등 인간의 특성을 파악하면서 착실하게 해결해 나가야 할 기초적인 검토 과제가 남겨져 있다.

둘째, 시스템 구성은 이용 형태에 따라 다르지만, 궁극적으로 휴먼 팩터를 기본으로 한 무안경식 입체 영상 디스플레이 시스템 검토가 중요하다. 특히, 다안식 시스템의 경우 레이저 등의 특수한 조명을 필요로 하지 말아야 하며, 알고리즘의 실시간 하드웨어화가 요구된다. 스테레오 영상은 한 시점에서만 입체감을 느낄 수 있기 때문에 다수의 관찰자가 허용되지 않으며, 관찰자의 움직임의 범위도 한정된다. 따라서 시점을 확장하고 각 시점 사이의 연속성을 보장하기 위해서는 디스플레이 단에서 중간 시점에 나타날 영상을 합성(intermediate view reconstruction : IVR)하여야 한다. 이와 관련하여 Anthony[32]는 변이 추출 방법에 따라 변이-보상 방식과 변이-비보상 전달 방식으로 나누어 변이를 추출하는 방법을 제시하기도 하였다. 다안식 입체영상 시스템에서는 이러한 중간영상 합성기술의 연구가 데이터량과 계산량 그리고 가려진 영역의 표현문제 등을 해결할 수 있을 거라 기대 되며, 지속적인 연구가 필요하다.

마지막으로 전송에 대한 문제인데, 이 부분에서는 디지털 화상 압축을 주축으로 한 알고리즘의 개발이 주요 과제가 될 것이다. 전송과 관련하여 무엇보다 중요한 점은 스테레오 영상에서의 변이 추출에 관한 문제이다. 경계 영역에 대한 표현문제나, 블록화 현상에 대한 연구가 더욱 필요할 것이며, 이와 관련하여 실시간 시스템에 기여 할 수 있도록 수행시간을 단축하는 알고리즘 개발이 필요하다. 또한 차 영상 부호화도 많은 진전이 필요하다. 현재까지

JPEG표준이나 DCT기반의 변형된 압축기법을 사용하는 경우가 대부분이지만, JPEG2000이나 MPEG-4에서 사용하는 웨이블릿을 이용한 기준 영상 및 차 영상 부호화도 선보여지고

있다. 이와 관련하여 차 영상의 특성을 고려한 압축방법과 낮은 비트율에서도 주관적인 화질의 개선과 압축효율이 좋은 알고리즘 연구가 필요하다.

* 본 연구는 과학기술부 국가지정 연구실 사업의 연구비 지원으로 이루어짐.

참고 문헌

- [1] Roel ter Horst, "A digital codec for 3D-TV transmission," International HDTV Workshop 92, Nov.1992, Kawasaki, Japan
- [2] M. G. Perkins, "Data compression of stereopairs," IEEE Trans. Communications, vol. 40, pp.684-696, Apr. 1992.
- [3] 김은수, 이승현 "3차원 영상의 기초" 技多利, pp.10-15, Nov.1998
- [4] Han-jo Jung, Eun-soo Kim "A study on the stereo image compression based on wavelet transform", 광운대학교 전자공학과 석사 졸업논문, 1999
- [5] J.F.McVeigh "Efficient Compression of Arbitrary Multi-view video signals", Carnegie Mellon Univ. Pittsburgh, Pennsylvania, 1996.
- [6] Michael G. Perkins, "Data Compression of Stereopairs" IEEE Trans, on comm.,vol.40, no.4 pp.684 - 696, Apr. 1992
- [7] P.Skerjanc and J. Liu, "A three camera approach for calculating disparity and synthesizing intermediate pictures," Signal Processing: Image Comm., vol.4, pp.55-64, 1991.
- [8] Jeffrey S. McVeigh, M.W. Siegel and A.G. Jordan, "Intermediate view synthesis considering occluded and ambiguously referenced image regions," Signal Processing: Image Comm., vol.9, no.1, pp.21-28, Nov. 1996.
- [9] 김경태, 박대원, 이천희, 오병주, "다시점 영상 세트의 데이터 압축을 위한 3-D구조 추정 방법과 MPEG-2 부호화 방법에 관한 연구", The 2nd International Workshop on 3-D Imaging Media Technology, pp.171-181, March, 1996.
- [10] Siegel, M. W., P. Gunatilake, S. Sethuraman, and A. G. Jordan, "Compression of stereo image pairs and streams," Proc. SPIE vol. 2177, pp. 258-268, February 1994.
- [11] Anthony Mancini, Disparity estimation and intermediate view reconstruction for noble applications in stereoscopic video, Master Thesis, McGill University, Feb. 1998.
- [12] Dimitrios Tzovaras, Nikos Grammalidis, and Michael G. Strintzis, "Object-based coding of stereo image sequences using joint 3-D motion/disparity compensation," IEEE Trans. On Circuits and systems for video tech., vol. 7, no. 2, April, 1997.

● 참고 문헌 ●

- [13] D.V. Papadimitriou and T.J. Dennis, "Epipolar line estimation and rectification for stereo image pairs," IEEE Trans. On image processing, vol.5, no.4, April 1996.
- [14] A. Sharaf, F.Marvasti, "Motion compensation using sparial transformations with forward mapping," Signal Processing: Image Comm., Vol.14, pp.209-227, 1999
- [15] A. Murat Tekalp, "Two-Dimensional Mesh-Based Visual-Object Representation for Interactive Synthetic/Natural Digital Video," Proceedings of the IEEE, vol. 86, no.6, June 1998
- [16] 김미현 "불규칙 삼각망을 이용한 3차원 영상의 순방향 변이 추정과 중간시점 영상 합성", 연세대학교 전자공학과 석사학위 논문, 1999.
- [17] Puri, R.V. Kollarits, B.G. Haskell, "Basics of stereoscopic video, now compression results with MPEG-2 and a proposal for MPEG-4," Signal Processing: Image Comm., vol.10, pp.201-234, 1997.
- [18] S. Sethurama, "Stereoscopic image sequence compression using multiresolution and quadtree decomposition based disparity and motion-adaptive segmentation" Ph.D. Thesis, Carnegie Mellon Univ., Pittsburgh PA, 1996.
- [19] D. Tzouvaras, M. G. Strintzis, and H. Sahinoglu, "Evaluation of multiresolution block matching techniques for motion and disparity estimation," Signal Processing: Image Communication, vol. 6, no.1, pp.59-67, Mar. 1994.
- [20] Won-Ho Kim, Jae-Young Ahn and Sung-Woong Ra, "An efficient disparity estimation algorithm for stereoscopic image compression", Consumer Electronics : IEEE Trans, vol.432, pp 165 -172, May 1997
- [21] W. E. L. Grimson, "Computational Experiments with a feature based stereo Algorithm", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 7, pp. 17-34, Jan. 1985.
- [22] J. E. W Mayhew and j. p. Fisby, "Psychophysical and computational studies toward depth perception", Artificial Intelligence vol. 17, pp.349-385, 1981.
- [23] Mi-Hyun Kim and Kwang-Hoon Sohn, "Edge-preserving directional regularization technique for disparity estimation of stereoscopic images," IEEE Trans. On Consumer electronics, vol.45, no.3, Aug. 1999.
- [24] Yao Wang, Ouseb Lee, "Use of 2-D deformable mesh structures for video coding, Part I-The synthesis problem: Mesh-based function approximation and mapping, " IEEE Trans. On Circuits and systems for video tech, vol.6, no. 6, Dec. 1996.
- [25] Yao Wang, Ouseb Lee, "Use of 2-D deformable mesh structures for video coding, Part II-The analysis problem and a region-based coder employing an active mesh representation," IEEE Trans. On Circuits and systems for video tech., vol. 6, Dec. 1996.
- [26] R.Korenen, F.Pereira, and L. Chiariglione, "MPEG-4: Context and Objectives," Signal Processing: Image Comm., vol.9, pp.295-304, May 1997.
- [27] Y. Nakaya and H. Harashima, "Motion compensation based on spatial transforms," IEEE Trans. Circuits and System for Video Technology, vol. 4, no. 3, pp.339-356, Jun. 1994.
- [28] W. Pennebaker, "JEPG technical specification, revision 8," Working Document No. JTC1/SC2/WG10/JPEG-8-R8, Aug. 1990.
- [29] Sethuraman, S., M. W. Siegel, and A. G. Jordan, "A Multiresolutional Region Based Segmentation Scheme for Stereoscopic Image Compression", Proc. SPIE Vol. 2419, pp.265-274, February 1995.
- [30] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on SPIHT" IEEE Trans circuits and systems for video Technology, vol 6, pp.243-250 June. 1996
- [31] A. N. Skodras, C. A. christopoulos and T. Ebrahim: "JPEG 2000: The upcoming still image Compression standard proceedings of the 11th portugness Conference on pattern Recognition , Porto, Portugal , may 11th-12th, pp.359-366, 2000
- [32] Anthony Mancini, "Disparity estimation and intermediate view reconstruction for noble applications in stereoscopic video", Master Thesis, McGill University, Feb. 1998.

필자소개



유 지 상

- 1985 서울대학교 전자공학과 졸업(공학사)
- 1987 서울대학교 대학원 전자공학과 졸업(공학석사)
- 1993 Purdue 대학교 전기공학과 졸업(Ph.D.)
- 현재 광운대학교 전자공학과 조교수



김 은 수

- 1978 연세대학교 전자공학과 졸업 (공학사)
- 1984 연세대학교 대학원 전자공학과 졸업 (Ph.D)
- 1987 CalTech 전기공학과 객원교수
- 현재 사단법인 3차원방송영상학회 (회장)
- 현재 광운대학교 전자공학과 /국가지정3차원영상미디어연구실(NRL) 교수



배 진 우

- 1998 순천대학교 전자공학과 졸업(공학사)
- 현재 광운대학교 전자공학과 석사과정



이 병 진

- 2001 순천대학교 전자공학과 졸업(공학사)
- 현재 광운대학교 전자공학과 석사과정