

네트워크에 적응적인 스케일러블 스트리밍 비디오 코딩을 위한 새로운 스캔 방법

(A New Scanning Method for Network-adaptive Scalable Streaming Video Coding)

박 광 훈 [†] 정 원 식 ^{**}

(Gwang Hoon Park) (Won-Sik Cheong)

요 약 본 논문에서는 MPEG-4 미세입자 계층 코딩 기법 등과 같은 망 상태 적응적인 계층적 스트리밍 코딩 기법에 적합한 새로운 스캔 방법을 소개한다. 제안된 스캔 방법은 사용자가 관심있는 영역의 영상정보 부호화를 최우선적으로 수행하여 전송하고 수신단에서도 해당 영상정보의 복호화를 최우선적으로 수행하도록 관리하여 복호된 비디오 영상의 주관적인 화질 향상을 보장하는 방법이다. 제안된 방법을 FGS 코딩 기법에 적용하면 특히 사용자가 관심 있는 영역의 주관적 화질을 약 1 dB ~ 3 dB정도 향상시킬 수 있다.

키워드 : 동영상코딩, MPEG, 미세입자 계층 코딩, 스케일러블 코딩, 스트리밍 비디오

Abstract This paper introduces a new scanning method for network-adaptive scalable streaming video coding methodologies such as the MPEG-4 Fine Granular Scalable (FGS) Coding. Proposed scanning method can guarantee the subjectively improved picture quality of the region of the interest in the decoded video by managing the image information of that interested region to be encoded and transmitted most-preferentially, and also to be decoded most-preferentially. Proposed scanning method can lead the FGS coding method to achieve improved picture quality, in about 1 dB ~ 3 dB better, especially on the region of interest.

Key words : MPEG, Fine Granular Scalable Coding, Scalable Coding, Streaming Video

1. 서 론

동영상 코딩방법에 있어 계층적 코딩(scalable coding)기법에 대한 요구는 크게 증가되고 있다. 특히 언제 어디서나, 누구에게나, 영상정보를 이용하여 통신이 이루어질 수 있도록 하는 이동통신 서비스나, 무선 인터넷 등의 등장으로, 노트북, 팜탑, PDA 등 다양한 컴퓨터와 결합된 정보가전을 이용하여 원격지의 영상정보를 얻고 관리하고 변경할 수 있기를 원한다. 그러므로 다양한 형태의 영상정보 가전들이 앞으로 출시될 것이며, 출시될 영상정보 가전제품의 복호화 능력이나 전송

환경 등은, 각각의 단말기가 가져야 하는 특성이나 적용 환경 때문에 자연히 다르게 된다. 이러한 문제를 해결하기 위하여, 영상을 코딩하여 전송하는데 있어서, MPEG-4에서는 계층적 코딩을 할 수 있도록 설계가 되어있다. 계층적 비디오 코딩은 수신단 측에서 낮은 품질에서 높은 품질까지 다양한 화질의 비디오를 수신할 수 있도록 엔코더측에서 계층적 비트스트림을 만들어서 전송하는 것이다. 계층적 코딩기법은 기본적으로 크게 기본계층(Base Layer)과 강화계층(Enhancement Layer)으로 구성된다. 송신단 측의 기본계층에서는 기본품질의 동영상 정보를 전송하며, 강화계층에서는 기본품질의 동영상에 더해져서 수신단 측에서 향상된 화질을 제공할 수 있는 영상정보를 부가하여 전송한다. 수신단 측에서는 단말기의 성능이나 전송선로의 상태에 따라 전송된 두 가지 계층의 영상정보를 선택하여 복호하게 된다. 즉, 복호기가 전송선로를 통하여 전달된 모든 정보에 대

[†] 종신회원 : 경희대학교 전자정보학부 교수
ghpark@khu.ac.kr

^{**} 정 회 원 : 한국전자통신연구원 방송미디어연구부 연구원
wscheong@etri.re.kr

논문접수 : 2002년 1월 8일

심사완료 : 2002년 4월 9일

한 복호 능력을 가지고 있지 않다면, 복호기는 최소 영상품질 보상계층인 기본계층에 대한 정보만을 복호화하고, 나머지 강화계층에 해당하는 정보는 복호하지 않고 버리게 된다. 고품질의 수신기의 경우에는 모든 계층의 정보를 다 받아들이고 고품질의 영상을 구현하게 된다. 이로써 고품질의 복호기를 소유한 사용자와 저품질의 복호기를 소유한 사용자 모두를 만족 시킬 수 있는 영상전송을 계층적 코딩기법을 이용하여 구현할 수 있다 [1,2,3,4].

기존의 계층적 코딩기법 등은, 전송선로의 상태가 급변하지 않고 상대적으로 안정된 경우에 적합하도록 설계되어 있다. 즉 수신단측에서 강화계층에서 송신된 비트스트림을 모두 수신하여야만 해당 영상 프레임의 복원할 수 있다. 만약 전송선로의 상태가 급변하여(전송선로가 수용 가능한 비트스트림의 전송대역폭(bandwidth)이 급변함: 인터넷 등과 같은 전송선로는 인터넷 사용자의 수 등과 같은 외부 요인에 의하여 각 사용자에게 할당 가능한 전송대역폭이 급변한다), 강화계층에 해당하는 비트스트림을 모두 전송 받지 못한 경우에는 해당 영상을 제대로 복원 할 수가 없고, 따라서 해당 영상 프레임의 전송을 포기하거나, 재전송을 송신단 측에 요구하거나 모든 비트스트림을 수신할 때까지 영상복원을 수행하지 못하거나, 또는 이전 프레임의 영상을 이용하여 전송 에러 숨김(error concealment)을 수행하여야 한다[1,2,5].

2. 미세입자 계층 코딩 기법

MPEG-4 미세입자 계층(Fine Granular Scalability: FGS) 코딩 기법은 계층적 코딩을 수행할 때, 송신단 측의 기본계층과 강화계층 엔코더에서 각각 코딩되어 전송된 비트스트림을 수신단 측에서 모두 수신하지 못한 경우에도, 그때까지 수신된 부분적인 비트스트림 만을 이용하여 전송된 영상의 복원이 가능하도록 하는 망 상태 적응적인 계층적 스트리밍 코딩(Network-adaptive scalable streaming coding) 기법으로, 기존의 계층적 코딩 방법이 안정된 전송선로를 고려하여 구현된 것의 단점을 보완할 수 있도록 설계되어 있다[1,2,6~11]. 이를 위하여 FGS 코딩 방법에서는, 송신단 측의 강화계층에서 향상된 화질의 영상 비트스트림을 구현하여 전송할 때, 비트플레인(Bit-plane) 단위로 전송하는 방법을 이용하여 수행한다. 즉 송신단 측에서 수신단 측으로 강화계층에 필요한 비트스트림을 전송할 때, 원 영상과 기본계층에서 전송되는 영상의 오차만을 전송하여 전송영상의 화질 향상을 이루도록 하는 것은 기존의 계층적 코

딩 방법과 비슷하지만, 강화계층에서 전송할 영상정보를 비트플레인 별로 나누어 가장 중요한 비트(Most significant bit: MSB)를 최우선적으로 전송하고, 그 다음 중요한 비트를 차례대로 전송하는 방식을 사용하여, 수신단 측에서 전송선로의 대역폭이 급변하여 영상복원에 필요한 모든 비트를 수신하지 못한 경우에도, 그때까지 수신된 비트스트림 만을 이용하여 전송된 영상의 복원을 어느 정도 수행할 수 있도록 하는 방법이다[6,9,10]. 예를 들어 25라는 영상정보를 전송한다고 가정하면, 이를 바이너리 숫자로 표현하면 "11001"이 되고 이는 5개의 비트플레인으로 구성되어 있다. 이 정보를 비트플레인별로 송신하기 위해서는 일단 송신단측에서 수신단측으로 모든 전송정보의 구성이 5개의 비트플레인으로 구성되어 있다고 알려주고, 가장 중요한 MSB부터 LSB(Least significant bit)까지 비트단위로 수신단으로 전송한다고 할 때, 첫번째 MSB가 전송완료 되었다면, 수신단측에서는 전송되는 정보가 16(10000) 이상의 값이라는 것을 알 수 있고, 두번째 비트열이 전송되었다면 24이상의 값(11000)이 전송될 것이라는 것을 알 수 있다. 만약 전송대역폭의 제약으로 더 이상의 비트스트림을 수신단측에 전송할 수 없다면, 수신단측에서 이제까지 전송된 비트열(11000)을 이용하여 24라는 숫자를 복원할 수 있다.

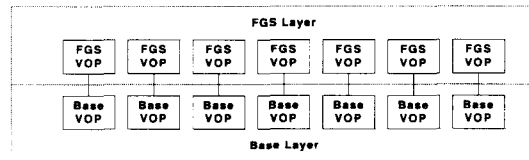


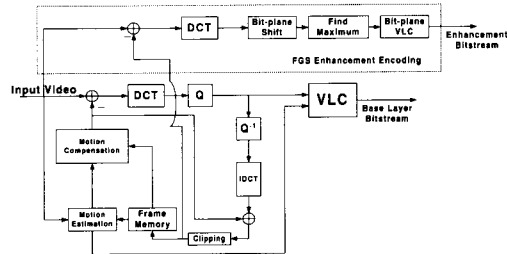
그림 1 기본적인 FGS 코딩 방법 구조

기본적인 FGS 코딩 방법의 구조를 살펴보면 그림 1과 같다[6]. 그림에서 보는 바와 같이 기본계층과 함께, 강화계층으로서 FGS 계층을 두고 있다. 기본 계층은 기존의 MPEG-4 비디오 코딩 기법을 그대로 사용하고 있다. 강화계층인 FGS 계층간에는, 계층간의 움직임 추정 및 보상 등 코딩효율을 증가시키는 방법론을 전혀 사용하지 않고 단지 기본 계층으로부터 코딩 효율을 증가시키는 방법만을 고려하고 있는 것이 특징이다[6].

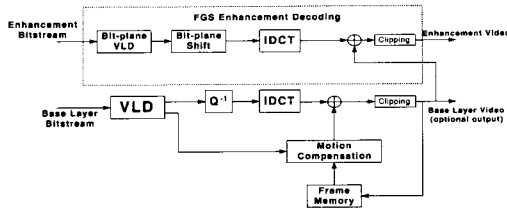
FGS 코딩 기법의 엔코더 구조는 그림 2(a)에 도시된 것과 같다. 그림에 도시된 것과 같이, 기본 계층은 MPEG-4 비디오 엔코딩 기법을 이용하고 있다[1,6]. 강화 계층의 FGS 엔코딩은 그림에서 도시된 것과 같이 원영상과 기본계층에서 비트스트림을 전송한 후 재생된

영상과의 오차(residues)를 구하는 단계와, DCT를 수행하는 단계, 비트플레인 쉬프트(Bit-plane shift) 단계, 최대값 찾기(Find Maximum) 단계, 그리고 비트플레인 별로 가변장 부호화 수행 (Bit-plane VLC) 단계로 나뉘어 수행된다.

오차를 구하는 단계에서는, 기본계층에서 코딩된 후 재생된 영상과(그림 2에서 Q^{-1} 와 IDCT를 통과한 후 Clipping 된 재생 영상) 원영상과의 차를 구하여 그 오차 값을 구하는 단계이다. 이산적 코사인 변환을 수행하는 단계에서는 앞에서 구해진 영상기반 오차 값을, 블록 단위(8×8 단위) DCT를 이용하여 DCT 도메인으로 변환하는 단계이다. 이때 선택적으로 좋은 화질을 갖는 블록이 필요하다면 해당 값을 최우선적으로 전송하여야만 하고 이를 위하여 비트플레인 쉬프트를 선택적으로 수행할 수 있다. 이를 선택적 강화(Selective Enhancement: SE)라고 정의하며 비트플레인 쉬프트 단계에서 수행한다[6,8]. 최대값 찾기 단계에서는 DCT를 수행한 모든 값들의 절대치에 대한 최대값을 구한다. 이 값은 해당 영상 프레임에 전송하기 위한 최대 비트플레인의 갯수를 구하는 데 사용된다. 비트플레인 가변장 부호화 단계에서는, 비트플레인 별로 블록 단위로 구해진 64개의 DCT 계수들은 (DCT계수의 해당 비트플레인의 비트:0 또는 1) 지그재그(zigzag) 스캔 순서로 하나의 행렬에 들어가게 되며, 각 행렬들은 VLC코드표에 따라 RLE(Run-length encoding)를 수행한다.



(a) 인코더단의 구조

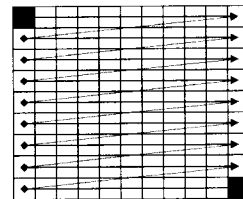


(b) 디코더단의 구조

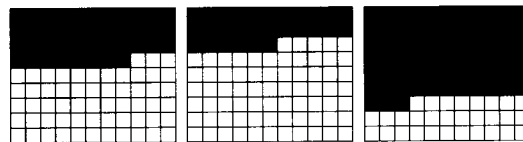
그림 2 FGS 코딩 기법의 구조

FGS 코딩 기법의 디코더 구조는 그림 2(b)에 도시된 것과 같이, 기본 계층에서는 MPEG-4 비디오 디코딩 기법을 그대로 사용하고 있다. FGS 강화 계층에서는 입력된 강화 비트스트림을 비트플레인 별로 가변길이 디코딩(VLD)을 수행한 후, 선택적으로 좋은 화질을 갖는 블록의 위치 등이 전송되었다면, 비트플레인 쉬프트를 선택적으로 수행한다. 비트플레인 별로 VLD가 수행되고 선택적으로 쉬프트가 수행되어 도출된 값에, 블록 단위의 IDCT를 수행하여 강화계층에서 전송된 영상을 복원하고 이를 기본 계층에서 복원된 영상과 합친 후에 해당 값들을 0과 255 값 사이로 클리핑하여, 최종적으로 향상된 화질의 영상을 복원하게 된다.

JPEG, H.263, MPEG등에 주로 사용되는 DCT를 이용한 동영상 코딩 기법들을 살펴보면, 매크로블록과 블록단위로 영상 데이터를 코딩하여 전송하는데 이때 모든 영상 프레임에 대한 엔코딩과 디코딩은 좌측상단(Upper-Left) 위치한 매크로블록부터 시작하여 우측하단(bottom-right) 부분의 매크로블록까지의 영상 데이터를 순차적으로 처리하는 레스터 스캔 방법(Raster Scan Method)을 사용하고 있으며, 이는 그림 3(a)에 도시되어 있다. 만약 이 스캔방법을 부분적인 비트스트림만을 수신하여도 영상복원이 가능하도록 하게 하는 FGS 코딩방법에 적용하게 되면 그림 3(b)에 도시된 것과 같이 상위부분의 부분적인 매크로블록이나 블록들만이 디코딩이 수행되어 그 복원 영상이 수신단 측 화면에 표시되게 된다. 이는 그림에서 검게 채워진 블록이 디코딩이 수행된 블록이며, 채워지지 않은 블록이 디코



(a) 레스터 스캔 방법



■ 디코딩이 수행된 매크로 블록 (검은 블록)
□ 디코딩이 수행되지 않은 매크로 블록 (흰색 블록)

(b) 레스터 스캔을 이용한 디코딩 수행방법

그림 3 레스터 스캔 방법

딩이 수행되지 않은 블록이다.

FGS코딩 기법에서는, 기본계층에서 전송된 비트스트림과 강화계층에서 부분적으로 수신되어 디코딩된 비트스트림이 합쳐져서 향상된 영상이 수신단에서 재생되는데, 그림 3(b)의 예와 같이 영상의 상위부분의 데이터만 강화계층에서 수신되어 디코딩이 완료되었다면, 재생 영상의 화질은 강화계층에서 디코딩이 완료된 부분의 화질만 향상되게 된다. 만약 이때 향상된 부분의 영상이 배경이나 주인공의 얼굴 등이 아닌 곳 등의 시청자가 주목하지 않는 곳이라면, 강화계층용 비트스트림을 수신하여 복원하는 단계가 사용자가 인지하지 못하는 부분을 향상하는 처리를 수행하여 쓸모 없이 처리된 것과 같게 된다.

3. 제안된 파문 스캔 방법

부분적인 비트스트림만을 수신하여도 영상복원이 가능하도록 하기 위해서는, 강화계층간에 영상의 코딩 효율을 증대시키는 강화계층간의 움직임 보상 등의 어떤 방법론도 사용하지 못하며, 단지 기본계층으로부터 코딩 효율을 극대화시키는 방법만을 사용하여야 한다. 그러므로 강화계층의 영상정보를 전송할 때 레스터 스캔 방법을 사용하여 코딩 할 필요 없이 임의의 스캔 방법을 사용하여 전송하여도 디코더단에서 전송된 비트스트림을 인과성(Causality)과 무관하게 예러없이 복원할 수 있다.

본 논문에서는 인간의 시각 시스템(Human Visual System: HVS)을 고려한 스캔 방법을 제안한다. 제안된 방법은 전송하여야 할 영상 프레임내의 임의의 특정 부분부터 우선적으로 엔코딩을 수행하여 수신단으로 전송한 후 수신단 측에서 우선적으로 전송된 부분부터 디코딩을 수행하도록 하여, 전송선로의 문제로 더 이상 수신할 수 없는 상황이 될 때, 그때까지 전송된 비트스트림만을 이용하여 전송된 영상의 복원을 하는 방법으로 향상된 화질을 제공하여야 하는 부분을 우선적으로 송/수신하는 스캔 방법이다.

제안된 스캔 방법의 구성은, 우선적으로 코딩하여 전송하여야 할 임의의 특정 위치가, 코딩하고자 하는 영상 프레임 내에서 결정되거나 되었다면, 해당 위치부터 코딩을 수행하여 점차 주변으로 코딩을 확장하여 수행하도록 한다. 즉 임의의 특정 위치의 영상정보를 최우선적으로 코딩한 후, 바로 이웃하는 외각에 위치하는 영상정보를 그 다음에 코딩하는 것을 반복하여 수행한다.

이는 호수면 위에 돌을 던졌을 때 돌이 수면 위에 떨어진 위치로부터 파문(Water Ring)이 일어나는 것과

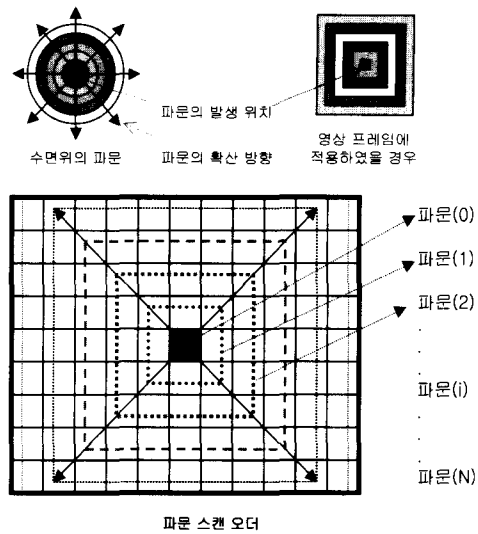


그림 4 파문 스캔 방법의 기본 원리

같은 이치로 기본 개념을 그림 4에 도시하였다. 이를 동 영상 코딩에 적용하였을 때, 파문발생 위치로부터 코딩을 시작하여 점차 주변으로 파문이 연쇄적으로 발생되면 해당위치의 데이터를 처리하는 것과 같다. 즉 파문발생 위치부터 시작하여 사각형 모양의 링을 계속적으로 그전에 처리한 사각형 모양의 링에 둘러싸는 형태의 스캔 방법을 의미한다. 그림에서 파문발생 위치(파문(0))의 데이터를 최우선적으로 처리한 후, 인접한 파문(1)의 데이터(파문(0)의 바로 외각에 위치한 데이터)를 처리하고, 파문(2) (파문(1)의 외각에 위치하는 16개의 인접한 데이터), 파문(3) 등에 위치한 데이터를 점차 파문이 확산

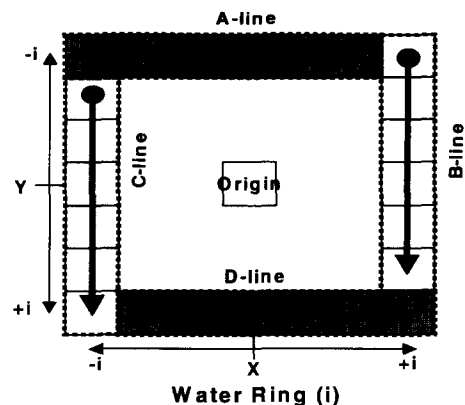


그림 5 i번째 파문의 발생위치

되는 형태로 연속적으로 처리하게 된다. 본 논문에서는, 제안된 방법이 파문이 확산되는 형태로 파문을 발생시키고, 해당 파문위치의 영상정보를 처리하는 스캔 방법이므로 파문스캔 방법이라고 명한다.

파문 스캔 방법에서 i 번째 발생된 파문의 데이터 위치를 그림 5에 도시하였다. 그림에서, 파문(i)는 A-line (Top line), B-line (right middle line), C-line (left middle line)과 D-line (bottom line)에 위치한 매크로블록으로 구성되어 있으며 파문 발생 위치로부터 i 번째 발생된 파문을 의미한다.

파문의 발생위치 결정과 해당 데이터 처리 방법은 다음과 같다.

스텝 1. 파문 발생위치(파문(0)) 결정 및 해당 데이터 처리

(a) 임의의 파문발생 위치를 정한다. 그림 5에서 "Origin"이라고 표시된 지점을 정한다. 이때 전송하고자 하는 영상 프레임의 중앙부분을 파문발생 위치로 정하거나 또는 사용자가 임의의 위치를 지정할 수 있다.

(b) 해당하는 파문발생 위치의 데이터를 처리한다. 엔코더에서는 엔코딩, 디코더에서는 디코딩을 수행한다.

스텝 2. 파문(i) 위치 결정 및 해당 데이터 처리

(a) 파문발생 위치로부터 i 번째 발생된 파문의 위치를 결정한다. 본 논문에서 파문(i): WaterRing(i)라고 정의한다.

(b) 파문(i)에 위치하는 데이터를 처리한다. 엔코더에서는 엔코딩, 디코더에서는 디코딩을 수행한다.

스텝 3. 영상 프레임 내의 모든 데이터가 처리될 때까지 스텝 2를 반복 수행한다.

파문 스캔 방법은 파문발생 위치 결정부와 파문발생부 그리고 파문의 위치에 해당하는 데이터를 처리하는 영상 코딩부로 크게 나눌 수 있다. 파문발생 위치 결정부에서는, 엔코더 측의 결정부에서 파문이 발생하는 임의의 위치를 결정하고 해당 좌표를 디코더 측으로 송신하는 것을 수행하는 부분이다. 디코더 측의 결정부에서는 엔코더 측에서 전송된 좌표를 바탕으로 영상 프레임 내의 파문 발생위치를 결정한다. 영상 프레임의 중앙부분 등의 파문발생 위치를 사전에 결정하여 엔코더와 디코더가 서로 약속하였다면 본 결정부의 역할을 생략한다. 파문 발생부에서는 앞서 설명한 파문발생 방법을 이용하여 i 번째 파문을 발생시키는 역할을 담당하며, 파문(i)의 발생위치를 영상 코딩부에 통보하여 영상 코딩부가 해당 영상 프레임의 코딩을 수행하도록 한다. 영상 코딩부에서는 파문 발생부에서 결정된 좌표에 해당하는 영상 데이터를 처리한다.

파문스캔 방법의 간단한 구현 예는 다음과 같다.

```

nRing: # of  $n^{\text{th}}$  Water Ring (Water Ring ( $n$ )) (initial value = 1)
nMB: # of MB's in each line (A, B, C, D) of Water Ring ( $n$ )
OriginX, OriginY: Origin coordinate ( $x$ ,  $y$ ).
CurrX, CurrY: Start (upper-left) coordinate of Water Ring ( $n$ ).
CodedMB( $x$ ,  $y$ ): Encoding or decoding function for designated MB( $x$ , $y$ ), in which VLC or VLD for each bit-plane are performed.
InBoundary( $x$ ,  $y$ ): A checking function to distinguish whether the designated MB( $x$ ,  $y$ ) is located at the outside of the picture.

Step 1: Coding of the 'Origin'
if(InBoundary(OriginX, OriginY))
CodedMB(OriginX, OriginY);

Step 2: Coding 'A-line' of Water Ring ( $n$ )
CurrX = OriginX - nRing;
CurrY = OriginY - nRing;
nMB = 2 * nRing;
for(  $j=0$ ;  $j < nMB$ ;  $j++$  )
if(InBoundary(CurrX+ $j$ , CurrY)) CodedMB(CurrX+ $j$ , CurrY);

Step 3: Coding 'B-line' of Water Ring ( $n$ )
 $x$  = CurrX + nMB;
for(  $j=0$ ;  $j < nMB$ ;  $j++$  )
if( InBoundary( $x$ , CurrY +  $j$ ) ) CodedMB( $x$ , CurrY +  $j$ );

Step 4: Coding 'C-line' of Water Ring ( $n$ )
for(  $j=1$ ;  $j \leq nMB$ ;  $j++$  )
if(InBoundary(CurrX, CurrY+ $j$ )) CodedMB(CurrX, CurrY+ $j$ );

Step 5: Coding 'D-line' of Water Ring ( $n$ )
 $y$  = CurrY + nMB;
for( $j=1$ ;  $j \leq nMB$ ;  $j++$ )
if(InBoundary(CurrX+ $j$ ,  $y$ )) CodedMB(CurrX +  $j$ ,  $y$ );

Step 6: if ( not Frame Coded )
{ nRing++; goto Step 2; }
else Stop;

```

위에서 제시한 파문스캔 알고리즘을 이용하여 영상정보 코딩시 CIF급 영상의 경우 각각의 비트플레인마다 총 396(22 X 18)개의 매크로블록 코딩 순서만 정하면 된다. 이는 엔코딩시 또는 디코딩시 소요되는 총 계산량에 비하여 무시할 정도로 미미한 계산량이며, 래스터 스캔 방법과 처리 속도를 비교하여도 파문 스캔시 추가로 소요되는 계산량은 무시할 정도이다.

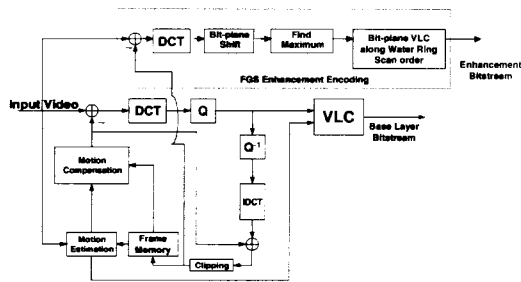
4. 파문 스캔 방법을 이용한 FGS 코딩 기법

파문 스캔 방법을 이용한 FGS 코딩 시스템의 엔코더 단 구조와 디코더단 구조를 그림 6에 도시하였다. 제안된 시스템에는, 각 비트플레인 별로 가변장 부호화

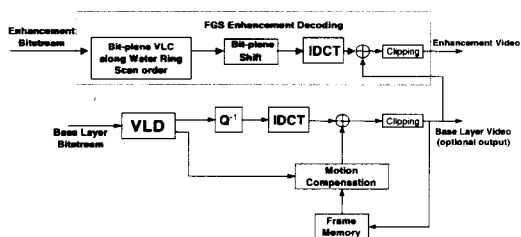
(Bit-plane VLC)를 수행할 때, 임의의 특정 위치로부터 파문 스캔을 수행하여 우선적으로 코딩할 매크로블록의 위치를 결정하고, 해당 매크로블록에 있는 영상정보를 위치 결정과 동시에 처리하는 단계가 엔코더단과 디코더단에 존재한다.

제안된 방법의 FGS 강화계층 엔코딩은, 그림 6(a)에 도시된 것과 같이 기존의 FGS 엔코딩 기법과 동일하지만, 파문스캔 방법에 따른 비트플레인 가변장 부호화 단계에서 각각의 비트플레인 별로 가변장 부호화를 수행할 때, 임의의 특정 위치로부터 파문스캔을 수행하여 우선적으로 코딩할 매크로블록의 위치를 결정하고, 결정된 코딩 순서에 따라 우선 순위에 맞게 블록 단위로 구해진 64개의 DCT계수들을 지그재그 스캔 순서로 하나의 행렬에 입력하고, 그들을 가변장 길이 코드표에 따라 RLE를 동시에 수행하는 것으로 변경되었다.

제안된 방법의 FGS 강화계층 디코딩은 그림 6(b)와 같이 엔코더와 역순으로 강화계층에 전송된 비트스트림의 디코딩을 수행하는데, 먼저 입력된 강화 비트스트림을 엔코더와 약속된(엔코더에서 시작 위치를 전송하여 수신된 위치나 또는 사전에 약속된 위치: 예를 들어 영상 프레임의 중앙 매크로블록) 파문발생 위치부터 비트플레인 별로 가변길이 디코딩(Bit-plane VLD)을 파문스캔 방법에 맞추어 수행하는 단계가 기존의 FGS코딩 방법에서 변경되었다.



(a) 엔코더 구조



(b) 디코더 구조

그림 6 파문 스캔 방법을 이용한 FGS 코딩 기법의 구조

5. 실험 결과

본 논문에서는 파문발생 위치를 엔코더와 디코더가 영상 프레임의 중앙부에 위치한 매크로블록에서 시작하는 것을 약속한 경우의 실험결과를 기존의 레스터 스캔 방법을 이용한 FGS 코딩 방법의 결과와 비교하였다.

그림 7에 파문 스캔 방법을 이용한 FGS 코딩 방법의 비교 실험 결과를 도시하였다. 그림에 표시된 2개의 영상은 MPEG-4 표준화 회의에서 주로 사용되는 '스테판(Stefan)' 영상 열(Sequence)을, CIF급으로 30Hz로 전송한다고 가정할 때, 기본계층은 256kbps(bit per second)로 전송하고, FGS 계층에서 강화 비트스트림을 엔코딩하여 전송하였지만 전송대역폭의 제약으로 최대 512kbps만의 비트스트림만을 디코더 단에서 수신 받아 재생된 영상을 보여주고 있다.

그림에서, ●로 표시된 영상은 기존의 레스터 스캔을 이용한 FGS 코딩 방법으로 재생된 영상이며, ●로 표시된 영상은 파문 스캔 방법을 이용한 FGS코딩을 수행한 결과 영상이다. '스테판' 인물(Tennis Player) 부분을 살펴보면 ●의 영상보다 ●의 영상에서, 보다 높은 주관적으로 향상된 화질을 관찰할 수 있다. 객관적인 화질을 PSNR(Peak Signal to Noise Ratio)을 이용하여 그래프로 표시하였는데, ●로 표시된 그래프는 Luminance (Y) 부분의 PSNR 비교 그래프이고 ●과 ●로 표시된 그래프는 각각 Chrominance(U, V) 부분의 PSNR 비교 그래프로서, 파문 스캔을 이용한 FGS 방법의 PSNR이 기존의 FGS방법에 비해 전반적으로 항상 약 2 ~ 3dB 가 높음을 알 수 있다. PSNR의 계산은 인간의 시각 시스템을 고려하여 CIF 영상의 중심부에 위치한 ROI (Region of Interest) 내의 PSNR을 계산하였으며, ROI는 14×10개의 매크로블록으로 구성되어 있다. 참고로 CIF영상에는 총 22×18개의 매크로블록이 있다. 대부분의 동영상 열을 분석해 보면, 동영상 촬영시 주요 물체의 표현 위치를 화면의 중앙부에 위치시키게 하고 있으며, 거의 대부분의 주요 물체는 본 논문에서 제시한 ROI영역에 위치하고 있다. '스테판'과 '포맨', '해안경비대', '카폰' 시퀀스의 ROI의 예를 그림 8에 도시하였다.

그림 7에서 실제 코딩 결과를 관찰한 바와 같이 주관적으로나 객관적으로도 파문 스캔 방법을 적용한 FGS 코딩 방법이, 기존의 FGS 방법보다 우수함을 관찰할 수 있는데, 화질차이가 나타나는 원인을 ●과 ●로 표기된 보조 그림(sub-figure)을 이용하여 분석할 수 있다.

그림에서, ●으로 표기된 보조 그림은 기존의 FGS 방법에서 비트플레인 별로 디코딩을 수행한 매크로블록을

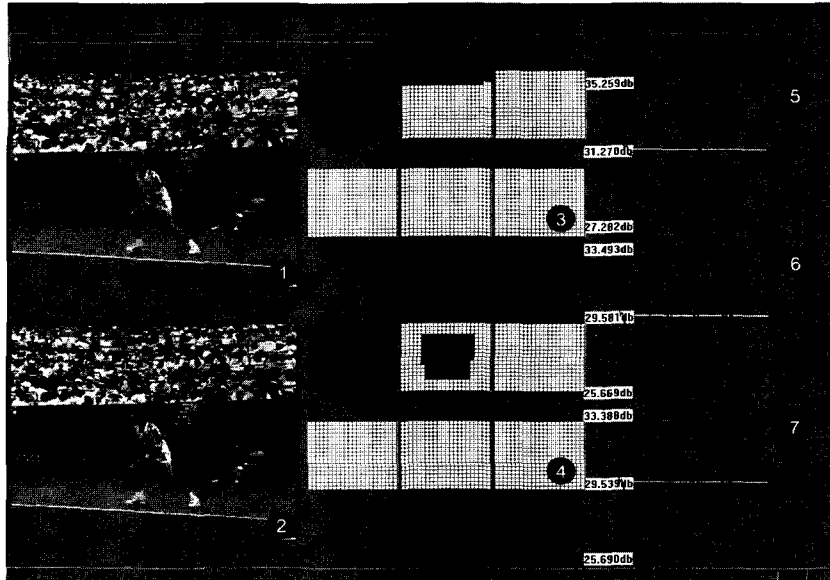


그림 7 MPEG-4 FGS 코딩 방법에 파문 스캔 방법을 적용한 결과

표현한 것으로, 전송대역폭의 제약으로 디코더 단에서 총 512kbps만을 수신하여 수신 완료된 영상정보만을 이용하여 디코딩을 수행한 결과 가장 중요한 비트플레인(●)의 그림에서 MSB라고 표시된 부분의 영상 데이터는 모두 코딩 완료되었지만(모두 검정색으로 채워져 있다), MSB-1 비트플레인의 데이터가 상위 1/4 이하만 디코딩이 완료된 상황으로(검정색으로 채워져 있지 않은 부분은 수신된 데이터가 없어서 디코딩을 수행하지 못한 경우임), 인간의 시각 시스템에 건주어 볼 때, '스테판' 영상 열의 인물에 해당하지 않는 상부 가장자리 부분만 화질 향상을 부여한 결과로서, 영상 프레임의 전체 화질을 주관적으로 평가할 때 화질이 상대적으로 저조함을 느끼게 된 것이다. 이에 비해 파문 스캔 방법을 이용한 FGS 코딩 방법은, ●로 표기된 부-그림에 디코딩이 완료된 매크로블록의 구성을 보여주고 있는데, MSB 비트플레인은 기존의 방법과 동일하게 모두 디코딩이 완료되었고, 역시 MSB-1 비트플레인의 영상 정보는 부분적으로 디코딩이 완료되었음을 보여주고 있다. 하지만 파문 스캔 방법을 이용하여 인간의 시각시스템에 적합하도록 영상 프레임의 중앙부부터 엔코딩과 디코딩을 수행하여 해당 위치의 영상정보를 최우선적으로 처리하여 그 결과를 보여 주고 있다. 그 결과, ●로 표시된 영상에서 살펴볼 수 있듯이 화면 중앙부의 영상의 화질이 결과적으로 우수함을 알 수 있다. 이는 인간의 시각 시스템에 적합하게 파문 스캔 방법을 적용하여, 영상 프레임의 중앙부분

부터 혹은 임의의 특정 위치부터 엔코딩을 수행하여 전송하고, 디코더 부에서도 영상의 중앙부분부터 혹은 임의의 특정 위치부터 디코딩을 수행하여 전송대역폭의 제약으로 더 이상 전송된 비트스트림을 모두 수신하지 못한 경우에도 영상 프레임의 중앙부분은 혹은 특정 위치는 항상 우수한 화질의 영상을 재생할 수 있도록 설계된 것이다.

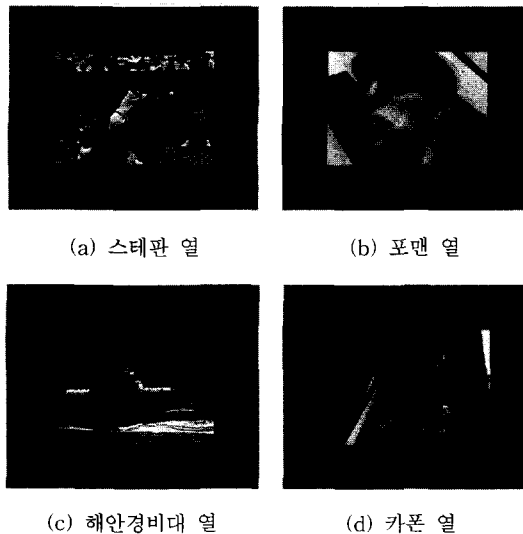


그림 8 ROI의 예

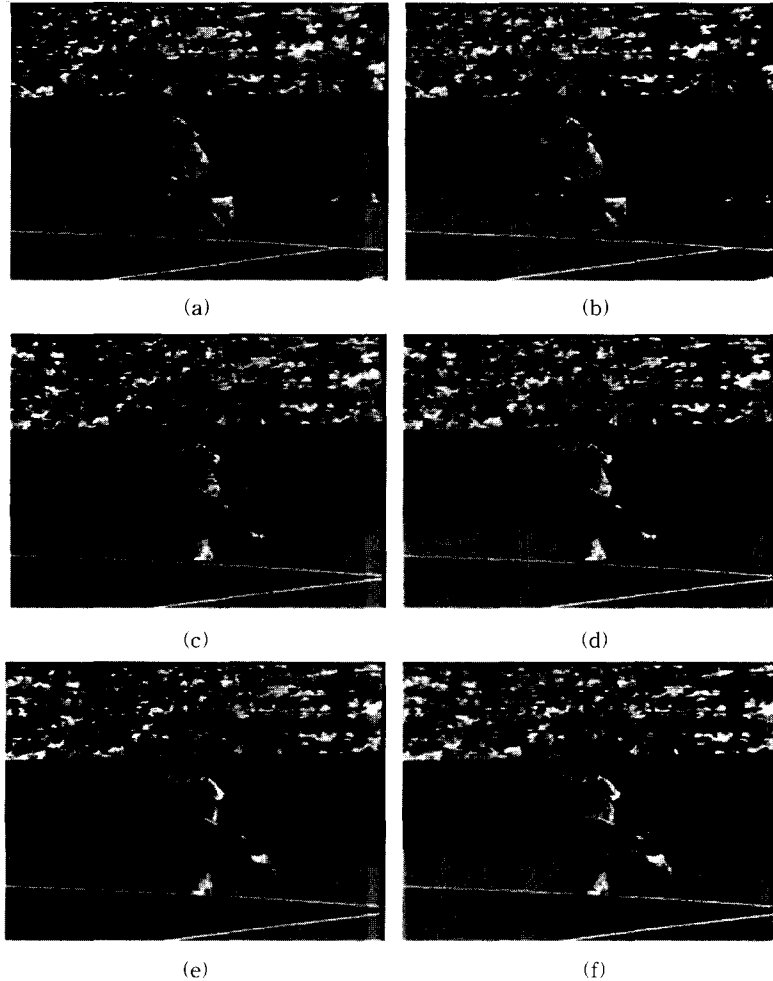


그림 9 FGS방법과 제안된 방법의 주관적인 화질 비교

그림 9에 그림 7에서 사용된 "Stefan" 영상 열 중에서 기존의 FGS와 파문 스캔을 이용한 FGS방법으로 재생된 결과 영상을 Snapshot으로 비교하였다. 그림에서 (a), (c)와 (e)는 FGS로 재생된 결과 영상이며 (b), (d)와 (f)는 제안된 방법을 이용하여 동일한 프레임 위치에서 재생된 결과 영상이다. 그림에서 주관적인 화질을 비교해 보아도, 제안된 방법의 결과가 상대적으로 우수함을 쉽게 발견할 수 있다.

그림 10은 '포맨' CIF급 10Hz 영상 열에 대해 64kbps의 비트스트림을 기본계층으로 전송하고 최대 512kbps까지의 수신 비트율에서, FGS 코딩과 파문 스캔을 이용한 FGS 코딩의 결과를, ROI내에서 구해진 PSNR(Y)을 이용하여 비교하였다. 사용된 ROI의 크기

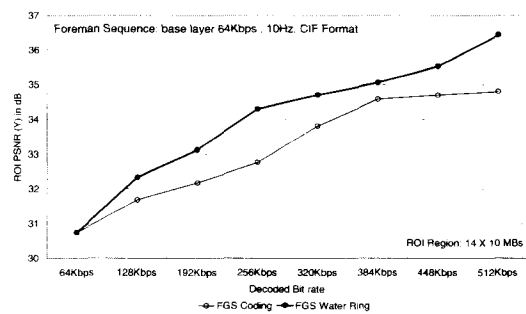


그림 10 '포맨' 영상 열에 대한 기존의 FGS 코딩, 파문 스캔을 이용한 FGS 코딩 방법에 대한 ROI내의 PSNR (Y) 성능 비교

는 프레임의 중심부에 위치한 14×10 개의 매크로블록으로 결정하였다. 그림에서 파문스캔을 이용한 FGS 코딩 방법이 ROI내에서는 기존의 FGS 코딩 방법의 PSNR(Y)을 1~2dB 정도 높여준다.

그림 11은 '스테판' CIF급 5Hz영상 열에 대하여 앞선 실험과 동일한 방법으로, 64kbps의 비트스트림을 기본계층으로 전송하고, 최대 512kbps까지의 수신 비트율에서, 기존의 FGS 코딩 방법과 파문 스캔을 이용한 FGS 코딩 방법의 성능을, ROI내에서 PSNR(Y)을 이용하여 비교하였다. 그림에서, ROI내의 PSNR을 살펴보면, 파문스캔을 이용한 FGS의 경우, 모든 수신 비트율에 대하여 FGS보다 약 2dB~3dB정도 우수한 결과를 얻어내었다. 이는 파문 스캔을 이용한 FGS 코딩 방법이 ROI내에서는, 기존의 FGS 코딩 결과를 매우 향상 시켜줄 수 있음을 보여주고 있다. ROI에서의 화질 향상은 전체 영역에 대한 객관적인 PSNR 수치에 대해서는 변화를 주지 못하나, 주관적인 판단에 의한 화질에 있어서는 매우 큰 차이를 갖게 된다.

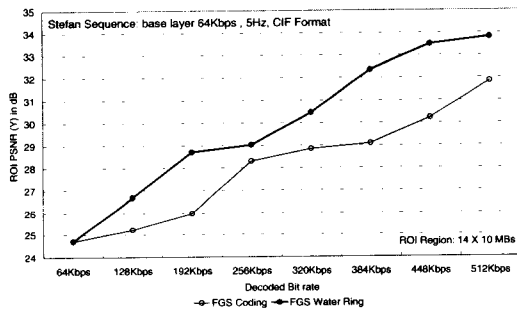


그림 11 '스테판' 영상 열에 대한 기존의 FGS 코딩, 파문 스캔을 이용한 FGS 코딩 방법에 대한 ROI의 PSNR (Y) 성능 비교

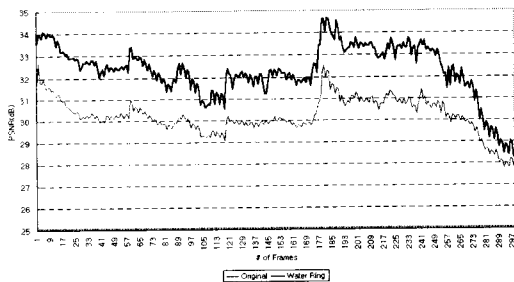


그림 12 '스테판' 영상열 (CIF급, 기본계층 비트율=256kbps, 최대 비트율=640kbps)에서, FGS 코딩 기법과 파문스캔을 이용한 FGS 코딩 기법의 PSNR(Y) 성능 비교

그림 12에 파문 스캔을 이용한 FGS 코딩 방법과 기존의 FGS 코딩 기법의 디코딩 영상의 화질을, 매 프레임마다 ROI에서 PSNR(Y)를 이용하여 비교한 그래프를 도시하였다. 비교 영상에는 CIF급 '스테판' 영상 열을 이용하였다. 그림에서 쉽게 알 수 있는 바와 같이 본 논문에서 제안하는 방법의 PSNR이 기존의 FGS 방법에 비하여 항상 1~3 dB정도 우수함을 보여주고 있다

6. 결론

본 논문에서는 인간의 시각 시스템에 적합하도록 시각적으로 중요한 특정 영상부분의 정보를 우선적으로 엔코딩을 수행하여 전송하고 수신단에서도 해당 특정부분의 영상정보를 우선적으로 디코딩 하게 하여 전송대역폭의 제약으로 더 이상 엔코더 단에서 전송된 모든 비트스트림을 디코더단에서 수신하지 못한 경우에도 특정한 중요부분의 화질을 항상 보장할 수 있게 하는 파문 스캔 방법을 제안하였다. 제안된 스캔 방법은 파문이 확산되는 형태로 파문을 발생시키고, 해당 파문위치의 영상정보를 처리하는 방법이다. 제안된 방법을 FGS 코딩 기법에 적용하면 사용자가 관심 있는 ROI내의 주관적 화질을 약 1dB~3dB정도 향상시킬 수 있다.

참고 문헌

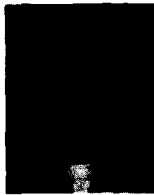
- [1] W. Li, J. Ohm, M. V. Schaar, H. Jiang, and S. Li, MPEG-4 Video verification model ver. 18.0, ISO/IEC/ JTC1/ SC29/ WG11/ N3908, 2001.
- [2] H. Radha, Y. Chen, K. Parthasarathy, and R. Cohen, "Scalable internet video using MPEG-4," Signal Processing: Image Communication, vol. 15, pp. 95~126, 1999.
- [3] W. Tan and A. Zakhor, "Real-time internet video using error resilient scalable compression and TCP-friendly transport protocol," IEEE Trans. Multimedia, vol. 1, pp. 172~186, 1999.
- [4] S. McCanne, M. Vetterli, and V. Jacobson, "Low-complexity video coding for receiver-driven layered multicast," IEEE J. Select. Areas Commun., vol. 16, pp. 983~1001, 1997.
- [5] 차호정, 오재학, "MPEG 스트리밍 시스템에서 QoS를 위한 프레임 드롭 필터링 시스템 구현", 한국정보과학회 논문지, vol. 26, no. 10, pp. 651~661, 1999.
- [6] MPEG-4 Fine granular Scalability Verification Model version 4.0, ISO/ IEC/ JTC1/ SC29/ WG11/ N3317, 2000.
- [7] Streaming video profile for FGS PDAM, ISO/ IEC/ JTC1/ SC29/ WG11/ M5673, 2000.
- [8] H. M. Radha, M. V. Schaar, and Y. Chen, "The

- MPEG-4 Fine-grained scalable video coding method for multimedia streaming over IP," IEEE Trans. On Multimedia, vol. 3, no. 1, pp. 53~ 68, March 2001.
- [9] MPEG-4 Fine granular Scalability Verification Model version 3.0, ISO/ IEC/ JTC1/ SC29/ WG11/ N3097, 1999.
- [10] MPEG-4 Fine granular Scalability Verification Model version 2.0, ISO/ IEC/ JTC1/ SC29/ WG11/ N2926, 1999.
- [11] 박광훈, "계층 부호화를 이용한 망 상태 적응적인 비디오 전송", Networked Video Workshop 2001, pp. 23~46, 2001.



박 광 훈

1985년 연세대학교 전자공학과(공학사), 1987년 연세대학교 전자공학과(공학석사). 1991년 Case Western Reserve Univ., Dept. of EEAP (M.S.). 1995년 Case Western Reserve Univ., Dept. of EEAP(Ph.D.). 1995년 ~ 1997년 현대전자 멀티미디어 연구소 책임연구원. 1997년 ~ 2001년 연세대학교 문리대학 전산학과 부교수. 2001년 ~ 현재 경북대학교 컴퓨터공학과 부교수. 관심분야는 멀티미디어, 비디오 신호처리, 패턴인식, 영상처리, 계산지능 등



정 원 식

1992년 경북대학교 전자공학과(공학사). 1994년 경북대학교 대학원 전자공학과(공학석사). 2000년 경북대학교 대학원 전자공학과 (공학박사). 2000년 5월 ~ 현재 한국전자통신연구원 방송미디어연구부 선임연구원. 관심분야는 영상신호처리, 영상 압축, 멀티미디어 시스템, remote sensing 등