DBpia



# 확장형 비디오 부호화(SVC)의 AR-FGS 기법에 대한 부호화 성능 개선 기법

종신회원 서 광 덕*, 정회원 정 순 흥**, 김 진 수***, 김 재 곤**

# Improved AR-FGS Coding Scheme for Scalable Video Coding

Kwang-deok Seo* ***Lifelong Member***,
Soon-heung Jung**, Jin-soo Kim***, Jae-gon Kim** ***Regular Members***

요 약

본 논문에서는 H.264의 확장형(scalable extension) 부호화 기법인 SVC(Scalable Video Coding)에서 채택하고 있는 AR-FGS(Adaptive Reference FGS) 기법의 재생화면 화질 향상을 위한 효과적인 방법을 제안한다. 표준 FGS(Fine Granularity Scalability) 기법에서는 FGS 계층의 부호화 성능 향상을 위하여 기본계층(base layer) 재생화면과 향상계층(enhancement layer) 참조화면에 대해 가중평균(weighted average)을 적용하여 FGS 부호화를 수행하는 AR-FGS 기법을 채택하고 있다. 그러나, 향상계층 부호화 정보가 비트스트림 절삭(bitstream truncation)에 의하여 FGS 복호기에 전달이 되지 못 할 경우 FGS 부호기와 복호기에 이용이 되는 참조화면의 차이로 인하여 움직임 보상 과정에서 오류의 전파(error drift)가 발생하여 FGS 계층에서 화질 저하를 초래하게 된다. 이를 해결하기 위하여 본 논문에서는 FGS 계층에서 움직임 보상에 이용될 예측신호를 구하기 위해 활용이 되는 향상계층 참조화면을 효과적으로 생성하기 위하여 사이클 블록 부호화(cyclical block coding)의 원리를 이용한다. 사이클 블록 부호화에서는 FGS 계층의 복호화 화질에 큰 영향을 미치는 중요 양자화 변환계수(quantized transform coefficient)를 초기 부호화 사이클에 포함시킴으로써 우선적으로 부호화 및 전송이 되게 하는 부호화 기술이다. 양자화 변환계수가 사이클 블록 부호화에 포함되는 순서가 앞설 경우 대역폭 감소로 인한 비트스트림 절삭이 적용될 때에도 복호기에 우선적으로 전달될 확률이 상대적으로 높다. 이러한 원리를 바탕으로 사이클 블록 부호화에서 각 사이클 별로 생성되는 비트스트림이 향상계층 참조화면의 생성에 기여하는 중요도에 따라 그 가중치를 다르게 조절함으로써 특정 부호화 사이클에서 생성된 비트스트림 정보가 절삭에 의해 FGS 복호기에 전달되지 못하더라도 복호화 시 그 영향을 최소화하여 화질 저하를 줄이는 방법을 제안한다. 제안된 방법을 이용하여 개선된 AR-FGS 기법을 구현할 경우 기존의 표준 방법에 비하여 재생화면의 화질이 최대 1dB 안팎으로 개선이 됨을 실험을 통해 확인하였다.

**Key Words** : SVC, FGS, AR-FGS

ABSTRACT

In this paper, we propose an efficient method for improving visual quality of AR-FGS (Adaptive Reference FGS) which is adopted as a key scheme for SVC (Scalable Video Coding) or H.264 scalable extension. The standard FGS (Fine Granularity Scalability) adopts AR-FGS that introduces temporal prediction into FGS layer by

* 연세대학교 컴퓨터정보통신공학부 (kdseo@yonsei.ac.kr), ** ETRI 방송미디어연구그룹
*** 한밭대학교 정보통신컴퓨터공학부
논문번호 : KICS2006-08-330, 접수일자 : 2006년 8월 1일, 최종논문접수일자 : 2006년 12월 14일



www.dbpia.co.kr

using a high quality reference signal which is constructed by the weighted average between the base layer reconstructed imageand enhancement reference to improve the coding efficiency in the FGS layer. However, when the enhancement stream is truncated at certain bitstream position in transmission, the rest of the data of the FGS layer will not be available at the FGS decoder. Thus the most noticeable problem of using the enhancement layer in prediction is the degraded visual quality caused by drifting because of the mismatch between the reference frame used by the FGS encoder and that by the decoder. To solve this problem, we exploit the principle of cyclical block coding that is used to encode quantized transform coefficients in a cyclical manner in the FGS layer. Encoding block coefficients in a cyclical manner places 'higher-value' bits earlier in the bitstream. The quantized transform coefficients included in the ealry coding cycle of cyclical block coding have higher probability to be correctly received and decoded than the others included in the later cycle of the cyclical block coding. Therefore, we can minimize visual quality degradation caused by bitstream truncation by adjusting weighting factor to control the contribution of the bitstream produced in each coding cycle of cyclical block coding when constructing the enhancement layer reference frame. It is shown by simulations that the improved AR-FGS scheme outperforms the standard AR-FGS by about 1 dB in maximum in the reconstructed visual quality.

## Ⅰ. 서 론

1990년대 초반부터 일반인에게 서서히 알려지게 된 인터넷은 최근에 백본망은 수십 기가비트의 대역폭을 처리하게 되었으며 가입자 망 또한 수십 메가비트를 지원할 수 있게 이르렀다. 이러한 통신인프라의 발전으로 인터넷의 초기에는 주로 텍스트와 정지 영상이 인터넷 트래픽의 대부분을 차지하였으나, 현재에는 음악이나, 동영상과 같은 실시간 멀티미디어 정보가 급격히 늘어나고 있으며, 이러한 추세는 더욱 가속화될 것으로 예측된다. 또한 WCDMA, HSDPA/HSUPA, CDMA2000 EVDO-Rev. A 등과 같은 3세대 및 3.5세대 무선 통신 서비스 뿐만아니라, 앞으로 다가올 4세대 무선 통신 서비스에서는 멀티미디어 정보 전송을 핵심 서비스로 정하고 있다. 특히, 동영상 전송 분야는 그 응용이 화상회의, 원격강의, 원격진료, VoD(Video on Demand) 등과 같이 매우 다양하고 유용하여 향후 큰 시장을 형성할 것으로 전망되며 관련 연구 또한 매우 활발한 상황이다. 그러나, 현재의 인터넷은 가입자 망의 종류가 다양하고 망의 대역폭이 시간에 따라 민감하게 변할 뿐만 아니라, 인터넷에 사용되는 프로토콜이 실시간 멀티미디어 정보 전달의 신뢰성을 보장해 주지 못하기 때문에 동영상 서비스의 확산이 어려운 상황이다.

인터넷 망이 실시간 동영상 서비스의 QoS(Quality of Service)를 보장해 줄 수 없기 때문에 영상 부호화 관점에서 이를 개선하기 위해 많은 연구들이 진행 되어 왔다. 이들 중 대표적인 부호화 기법 중의 하나가 계층부호화(layered coding) 기술이다[2].

계층부호화는 부호화된 원래의 단일 스트림을 영상 정보의 중요도를 기준으로 차별화된 화질과 전송 우선 순위를 갖는 계층화된 스트림 구조로 만드는 방법이다. 송신측은 주어진 대역폭에서 가능한 만큼의 계층화된 스트림 정보만을 전송함으로써 망의 대역폭 변화에 어느 정도 유연하게 대처할 수 있다. 그러나, MPEG-2와 H.263에서 표준으로 채택한 계층부호화 기술은 동일한 비트율을 적용한 단일계층(single layer) 부호화 방법에 비하여 화질이 많이 열화되는 단점이 있고, 계층의 개수를 필요한 만큼 많이 만들 수 없으므로 망의 동적인 대역폭 변화에 유연하게 대처할 수 없다는 단점이 있다. 이러한 기존의 계층부호화 기술의 한계와 단점을 극복하기 위해 MPEG-4에서 새롭게 제안된 기법이 비트평면(bit-plane) 부호화 기반의 FGS(Fine Granularity Scalability) 기술이다[1],[3],[9]. FGS는 부호화된 스트림의 전송 비트율을 망에서 가용한 대역폭에 간단하면서 정확하게 맞추는 것을 목표로 하며 1바이트 단위의 정밀도로 목표 전송율을 만족시킬 수 있다. 그러나, MPEG-4 FGS는 부호화 효율 측면에서 MPEG-4의 베이스라인 프로파일(baseline profile)에 비하여 성능이 많이 떨어지는 것으로 판명이 되었고, 이 문제를 개선시키기 위한 연구결과들이 다수 발표되었지만[12-14], 부호화 효율이 크게 개선되지는 못했다. 따라서, MPEG-4 FGS는 표준화 완료 이후에 상용 서비스에 활용이 되지 못하고 있는 실정이다[10],[11].

이러한 MPEG-4 FGS의 문제점을 근본적으로 해결



www.dbpia.co.kr

하기 위해 개발된 기술이 H.264 기반의 SVC(Scalable Video Coding)에 포함되어 있는 새로운 FGS 기술이다. SVC는 계층화된 하나의 비디오 스트림을 이용하여 다양한 전송 네트워크와 수신 단말에 적응적인 서비스를 가능하게 하는 확장형 비디오 부호화 방법으로 공간적(spatial), 시간적(temporal), 및 화질(SNR)에 대한 확장성 기능을 제공한다[4]. SVC의 SNR 확장성을 구현하기 위하여 새로운 FGS 기술이 개발 되었는데, 이 FGS 기술에는 압축 데이터의 중요도에 따른 차별적인 전송을 위하여 기존의 MPEG-4 FGS에서 채택한 비트평면 부호화(bit-plane coding) 대신에 사이클 블록 부호화(cyclical block coding) 를 채택하고 있다[8].

SVC의 FGS에서는 FGS계층의 부호화를 위하여 기본적으로 두 가지 부호화 방식을 채택하고 있다. 그림 1과 2는 FGS계층이 1개인 경우 이 두 가지 부호화 방식의 구조를 나타낸다. 그림 1은 FGS 계층에서 수행되는 움직임 예측부호화(motion compensated coding)에 기본계층에서 재생된 화면만을 참조화면으로 이용하는 기본계층 정보 기반의 FGS 부호화 방법의 구조를 나타낸다. 이 구조의 특징은 FGS 계층의 움직임 예측부호화가 향상계층 정보와 독립적으로 수행이 된다는 것이다. 향상계층 정보의 경우 망의 가용 대역폭이 감소할 경우 비트스트림 절삭(bitstream truncation)에 의하여 가용 대역폭을 초과하는 일부 비트스트림 정보는 수신측에 전송이 되지 않는다. 따라서, 그림 1의 구조에서는 일부 향상계층의 비트스트림 정보가 비트스트림 절삭에 의해 수신측에 전달이 되지 못하더라도, FGS 계층에서의 움직임 예측부호화가 향상계층 정보와는 독립적으로 수행이 되므로 FGS 계층의 화면을 오류 없이 복원할 수 있다. 또한, FGS 부호기와 복호기의 움직임 예측부호화에서 각각 사용이 되는 참조화면 간에 불일치가 발생하지 않으므로 여러 장의 화면을 통한 오류 전파(error propagation) 현상을 차단할 수 있다. 그러나, 그림 1의 구조에서는 FGS계층에서의 움직임 예측부호화에 기본계층으로부터 얻어진 저화질의 재생화면 만을 사용하게 되므로 예측의 정확성이 떨어져서FGS 계층의 부호화 효율이 급격히 저하되는 문제점이 있다.

그림 2는 그림 1의 단점을 보완하기 위한 표준 AR-FGS(Adaptive Reference FGS)의 기본 구조로서, 향상계층에서 재생된 고화질의 화면을 FGS 부호기의 움직임 예측부호화를 위한 참조화면의 일부로서 활용하고 있다[5],[6]. 여기서, '일부로서 활용한다'는 의미는 FGS계층에서 예측부호화를 위한 예측신호(predictor signal) 생성에 기본계층에서 재생된 저화질의 참조화면 만을 전적으로 활용하지 않고, 이 기본계층의 참조화면과 직전 화면의 향상계층에서 재생된 고화질의 참조화면에 대해 가중평균(weighted average)을 적용함으로써 새로운 예측신호를 생성하여 활용한다는 것이다. 그림 2에서 음영으로 처리된 박스에 보이듯이 n번째 화면의 FGS 계층을 부호화 하기 위하여 α, β의 가중치(weighting factor)를 이용하여 가중평균을 적용하는 이유는 비트스트림 절삭에 의해 향상계층 정보를 복호기가 수신하지 못하는 경우를 대비하여, 향상계층에서 생성된 참조화면에 대한 움직임 예측 부호화의 의존도를 조절하기 위함이다. 즉, α, β 를 모두 1보다 작은 값으로 설정함으로써, FGS 부호기와 복호기의 예측 부호화 과정에서 이용이 되는 참조화면 간의 불일치로 인하여 발생하는 오류 신호의 크기를 감쇄시키며, α, β 값을 0에 가깝게 설정할수록 이 오류 신호를 더욱 빠른 속도로 감쇄시킬 수 있다.

그러나, AR-FGS의 이러한 접근방법은 오류 전파 문제의 근본적인 해결책이 될 수는 없기 때문에 본 논문에서는 이 문제를 보다 효과적으로 해결하기 위한 방법을 제안한다. 본 논문에서는 FGS 계층의 움직임 예측부호화에 부분적으로 이용이 되는 향상계층에서 생성되는 참조화면을 효과적으로 생성하기 위하여 사이클 블록 부호화(cyclical block coding)의 기본 원리를 이용한다. 사이클 블록 부호화는 SVC의 표준 FGS 및 AR-FGS에서 채택하고 있는 스캔 방법으로서 각 4x4 블록마다 지그재그 스캔 순서에 의해 얻어진(*EOB*, *Run*, *Level*) 심볼을 중요스캔(significant scan)과 정제화스캔(refinement scan) 대상으로 구분하여 중요스캔에 포함되는 심볼을 초기 부호화 사이클에 포함시키고, 정제화 스캔에 포함되는 심볼을 후반부 사이클에 포함시킴으로써, 중요 정보를 먼저 전송하는 원리를 따른다[4],[7]. FGS 계층의 양자화 변환계수(quantized transform coefficient) 가 사이클 블록 부호화에 포함되는 사이클 순서가 앞 부분일 경우 대역폭 감소로 인하여 비트스트림 절삭이 적용될 때에도 복호기에 우선적으로 전달될 확률이 높다. 이러한 원리를 바탕으로 사이클 블록 부호화에서 각 사이클 별로 생성되는 비트스트림이 향상계층 참조화면 생성에 기여하는 중요도에 따라 가중치(weighting factor)를 조절함으로써 특정 사이클에서 생성된 비트스트림 정보가



www.dbpia.co.kr

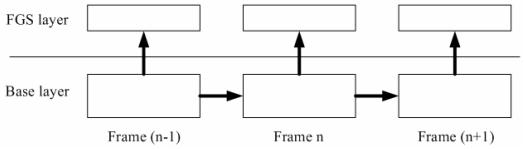



그림 1. 기본계층 정보 기반의 표준 FGS 부호화 구조.

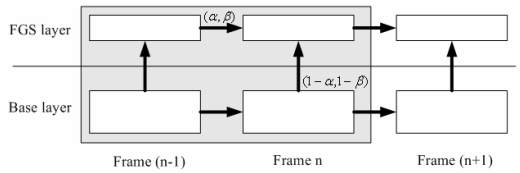



그림 2. 기본계층+향상계층 정보 기반의 AR-FGS 부호화 구조.

절삭에 의해 FGS 복호기에 전달되지 못 하더라도 복호화 시 그 영향을 최소화하여 화질 저하를 줄일 수 있는 방법을 제안하고자 한다.

본 논문의 구성은 다음과 같다. II장에서 표준 AR-FGS 부호화 방법에 대해 설명하며, III장에서 사이클 블록 부호화의 원리를 기반으로 AR-FGS 부호화의 성능을 향상시키는 방법을 제안한다. IV장에서 제안된 방법의 개선효과를 실험결과를 통해 검증하며 V장에서 결론을 맺는다.

## Ⅱ. 표준 AR-FGS 부호화 기법

SVC 표준화를 위하여 활동중인 JVT(Joint Video Team)에 제출된 AR-FGS 기술에 관한 표준화 기고문(contribution document)은 [5]이며, AR-FGS는 JVT에서 기술적인 검토를 통해 가장 최근에 출판된 SVC 표준 문서 JSVM 6.0[6]에 표준기술로서 포함이 되었다.

그림 3은 [5]와 [6]에도 제시되어 있는데, 그림 2의 AR-FGS 부호화 구조에서 n번째 화면의 FGS 계층을 부호화하기 위해서(n-1)번째 화면의 향상계층 참조화면과 n번째 화면의 기본계층 재생화면을 동시에 이용하는 AR-FGS의 세부 구현 원리를 나타내고 있다. 그림 3에서 n번째 화면의 FGS 계층에서 부호화될 $M \times N$ 크기의 블록인 $X^n$을 AR-FGS 기법에 의하여 부호화 하기 위해서 기본 계층의 재생화면 뿐만아니라 향상계층에서 얻어지는 참조화면을 동시에 이용하여 예측신호 $R_a^n$을 구한다. 즉, $R_a^n$은 기본계층에서의 재생신호인 $X_b^n$과 향상계층에서의 참조신호인 $R_e^{n-1}$ 간의 가중평균에 의해 얻어진다. $X_b^n$은 $X^n$과 화면상에서 동일한 블록 위치에 해당하며 기본계층에서 재생된 저화질의 블록을 의미한다. $R_e^{n-1}$은 (n-1)번째 화면의 향상계층에서 재생된 고화질의 참조화면으로부터 움직임 추정에 의해 추출되며 움직임 보상을 위한 참조블록을 나타내며, 움직임 추정에 사용되는 움직임 벡터는 $X_b^n$을 예측 부호화하기 위하여 기본계층에서 사용한 움직임 벡터와 동일하다. 즉, $R_e^{n-1}$은 $X^n$과 가장 유사한 블록을 기본계층에서 사용된 움직임 벡터를 이용하여 향상계층에서 재생된 고화질의 참조화면으로부터 추출한 신호를 의미한다. AR-FGS에서는 $X_b^n$과 $R_e^{n-1}$에 대해 가중평균을 적용하여 $R_a^n$을 얻는데, $X_b^n$에 대해 움직임 보상을 적용한 후에 얻게 되는 차분신호(differential signal)에 정수변환(integer transform) 및 양자화를 순차적으로 적용하여 생성되는 블록인 $Q_b^n(u,v)$의 계수들이 어떠한 값을 갖는지에 따라서 가중평균에 이용할 가중치를 다르게 적용한다[5],[6]. $Q_b^n(u,v)$은 다음 식과 같이 표현될 수 있다.

$$Q_b^n(u,v) = Quant\{IntTransf(X_b^n - R_b^{n-1})\} \quad (1)$$

이 식에서 $R_b^{n-1}$은 $X_b^n$에 대한 움직임추정에 의해 (n-1)번째 화면의 기본계층에서 재생된 참조화면으로부터 추출된 예측부호화 블록을 의미한다. 움직임 보상을 통해 $X_b^n$과 $R_b^{n-1}$간의 차분신호가 얻어지고 그 결과에 차례대로 정수변환(*IntTransf*)과 양자화(*Quant*)를 적용한 신호가 결국 식 (1)에 정의되어 있는 $Q_b^n(u,v)$이다.

만약 $Q_b^n(u,v)$의 모든 $(u,v)$ 주파수 위치에서의 계수값(coefficient)이 0일 경우 $R_a^n$은 다음 식과 같은 가중평균에 의해 구해진다.

$$R_a^n = (1-\alpha)\cdot X_b^n + \alpha \cdot R_e^{n-1}, \text{ if } Q_b^n(u,v) = 0 \text{ for all } (u,v). \quad (2)$$

식 (2)에서 $\alpha$는 그림 2에 표시된 가중치를 의미하며 그 값의 범위는 $0 \le \alpha \le 1$ 이다.

만약, $Q_b^n(u,v)$의 모든 계수의 값이 0이 아닐 경우에는 $Q_b^n(u,v)$의 각 $(u,v)$주파수 위치에서의 계수



www.dbpia.co.kr

의 값에 따라서 $R_a^n$ 은 다음의 식을 통해 얻어진다.

$$F_{R_a^n}(u,v) = (1-\beta)\cdot F_{X_b^n}(u,v) + \beta \cdot F_{R_e^{n-1}}(u,v), \quad \text{for } (u,v) \text{ with } Q_b^n(u,v) = 0. \tag{3}$$

$$F_{R_a^n}(u,v) = F_{X_b^n}(u,v), \quad \text{for } (u,v) \text{ with } Q_b^n(u,v) \neq 0. \tag{4}$$

식 (3)과 식 (4)에서 $F_{X_b^n}$ 과 $F_{R_e^{n-1}}$ 은 각각 $X_b^n$ 과 $R_e^{n-1}$ 에 대해 정수변환이 적용된 블록이며 $F_{X_b^n} = f(X_b^n)$ 와 $F_{R_e^{n-1}} = f(R_e^{n-1})$ 로 표현된다. $\beta$ 는 그림 2에 표시된 가중치를 의미하며 그 값의 범위는 $0 \le \beta \le 1$ 이다. 화소영역 신호인 $R_a^n$ 은 식 (3) 또는 식 (4)를 계산한 후에 얻어진 $F_{R_a^n}$ 에 대해 역변환(inverse transform)을 적용함으로써 구해낼 수 있다.

한편, 식 (2)의 경우 모든 $(u,v)$ 에 대하여 $Q_b^n(u,v) = 0$ 이므로 식 (1)에 $Q_b^n(u,v) = 0$ 를 대입할 경우 다음의 관계식을 유도할 수 있다.

$$X_b^n = R_b^{n-1} \tag{5}$$

식 (5)의 결과로부터 식 (2)와 식 (3)을 각각 다음 식과 같이 표현할 수 있다.

$$R_a^n = R_b^{n-1} + \alpha \cdot (R_e^{n-1} - R_b^{n-1}), \quad \text{if } Q_b^n(u,v) = 0 \text{ for all } (u,v). \tag{6}$$

$$F_{R_a^n}(u,v) = F_{R_b^{n-1}}(u,v) + \beta \cdot (F_{R_e^{n-1}}(u,v) - F_{R_b^{n-1}}(u,v)), \quad \text{for } (u,v) \text{ with } Q_b^n(u,v) = 0. \tag{7}$$

식 (6)에서 $R_d^{n-1} = R_e^{n-1} - R_b^{n-1}$ 로 정의하고 $R_d^{n-1}{}' = \alpha \cdot R_d^{n-1}$ 로 표시한다면, 식 (5)의 조건으로부터 식 (6)은 다음과 같이 표현된다.

$$R_a^n = X_b^n + R_d^{n-1}{}' \quad \text{if } Q_b^n(u,v) = 0 \text{ for all } (u,v). \tag{8}$$

유사한 방법에 의해 식 (7)에서 $F_{R_d^{n-1}}(u,v) = F_{R_e^{n-1}}(u,v) - F_{R_b^{n-1}}(u,v)$ 로 정의하고 $F_{R_d^{n-1}}{}'(u,v) = \beta \cdot F_{R_d^{n-1}}(u,v)$ 로 표시한다면, 식 (7)은 다음과 같이 표현된다.

$$F_{R_a^n}(u,v) = F_{X_b^n}(u,v) + F_{R_d^{n-1}}{}'(u,v), \quad \text{for } (u,v) \text{ with } Q_b^n(u,v) = 0. \tag{9}$$

한편, 식 (4)와 식 (9)를 비교해 보면 $Q_b^n(u,v)$ 의 값에 따라서 $\boldsymbol{R_d^{n-1}}$ 에 대한 변환이 적용되는지의 유무에 그 차이가 있으며, 식 (4)와 식 (9)에서 사용되는 $F_{R_d^{n-1}}{}'(u,v)$ 는 각각 다음과 같이 표현될 수 있다.

$$F_{R_d^{n-1}}{}'(u,v) = \begin{cases} 0 & \text{for } (u,v) \text{ with } Q_b^n(u,v) \neq 0, \\ \beta \cdot F_{R_d^{n-1}}(u,v) & \text{for } (u,v) \text{ with } Q_b^n(u,v) = 0. \end{cases} \tag{10}$$

식 (9)와 식 (10)으로부터 식 (4)는 $\beta = 0$ 이 되는 특수한 경우임을 알 수 있고, 이는 향상계층 정보에 대한 가중치를 0으로 설정하여 FGS 계층에서의 예측부호화가 전적으로 기본계층 정보로부터 이루어짐을 의미한다. 즉, 식 (4)에서는 향상계층 정보를 FGS계층의 예측 부호화에 활용하는 AR-FGS의 기본적 원리가 적용이 되지 않으며, 그림 2의 기본계층 정보 기반의 표준 FGS의 부호화 원리를 적용되는 것과 동일한 효과를 나타낸다.

결과적으로 $Q_b^n(u,v)$ 의 계수에 따라 식 (4), (8), (9)에 의해 계산되어지는 $R_a^n$ 은 그림 3과 같은 부호화 구조를 통해 실제적으로 구현이 된다. 이상의 과정을 통해 얻어진 $R_a^n$ 을 예측신호로 이용하여 $X^n$

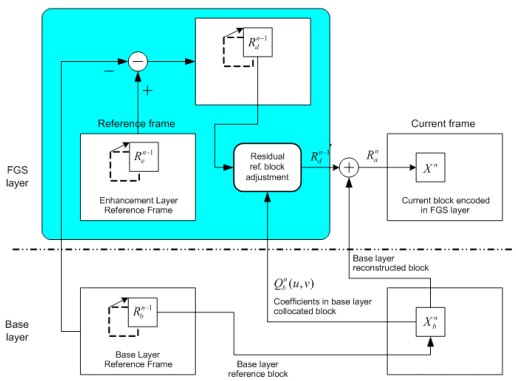



그림 3. AR-FGS 부호화 구현 원리[5],[6].



www.dbpia.co.kr

에 대한 움직임보상을 수행하게 되면 $X^n$과 $R_a^n$간의 차분신호를 얻게 되고, 이렇게 얻어진 차분신호에 정수변환 및 양자화를 차례대로 적용하여 부호화하게 된다. 한편, FGS계층에서 사용되는 양자화 파라미터(quantization parameter)는 그 크기가 기본계층에서 사용된 양자화 파라미터보다 6만큼 작은 값을 사용하여 기본계층보다 더 정밀한 양자화가 적용이 된다.

## Ⅲ. 개선된 AR-FGS 부호화 기법

그림 4는 표준 AR-FGS 부호화 기법의 성능을 개선시키기 위하여 본 논문에서 제안하는 개선된 AR-FGS 부호기의 구조이다. 기존의 AR-FGS부호기와 비교했을 때 크게 달라지는 부분은 음영으로 표시된 'Cycle Adaptive Coefficient Reconstruction' 블록인데, $\lambda$라는 새로운 가중치(weighting factor)를 도입하게 된다. 그림 4는 그림 3의 구현 원리에 의해 예측 신호 $R_a^n$을 얻은 후의 부호화 과정을 세부적으로 나타내고 있다. 그림 4에서 $R_a^n$은 II장에서 설명된 원리에 의해 식 (8) 또는 식 (4)와 식 (9)를 통해 얻어지는데, 'Predictor signal generator' 블록이 이 과정을 수행하게 된다.

$R_a^n$을 예측신호로 이용하여 $X^n$에 대한 움직임보상을 수행하게 되면 $X^n$과 $R_a^n$간의 차분신호 $E^n$을 얻게 되며, $E^n$은 4x4정수변환(integer transform)에 의해 $\hat{E}^n$로 변환이 된다. 사이클 블록 부호화(cyclical block coding) 블록에서는 $\hat{E}^n$의 변환 계수들을 지그재그 스캔에 의해(*EOB*, *Run*, *Level*) 심볼로 정렬한 후, 사이클 블록 부호화에 의해 부호화 사이클

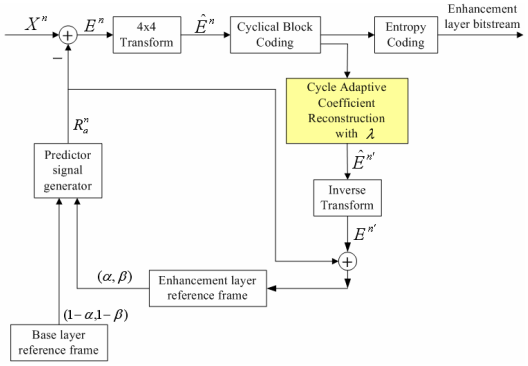



그림 4. 제안된 AR-FGS 부호기 구조.

별로 순서를 정하여 심볼들을 엔트로피 부호화 하게 된다[7],[8]. 그림 5는 3개의 변환계수(transform coefficient) 만을 갖는 4개의 변환블록(Block 0, Block 1, Block 2, Block 3)에 대한 사이클 블록 부호화의 적용 예를 보여준다. 이 예에서는 정제화 스캔(refinement scan) 대상이 되는 변환계수는 존재하지 않으며 중요스캔(significant scan)에 포함이 되는 변환계수만 있는 경우를 나타내고 있다. 그림에서 첫번째 부호화 사이클(coding cycle 0)에서 Block 0의 AC1, Block 1의 DC0, Block 2의 DC0, Block 3의 AC1이 중요스캔에 포함되어 차례대로 부호화 된다. 두번째 부호화 사이클(coding cycle 1)에서는 Block 0의 AC2, Block 1의 AC2, Blcok 2의 AC1, Block 3의 AC2가 중요스캔에 포함되어 부호화 된다. 마지막 부호화 사이클(coding cycle 2)에서 Block 0, Block 1, Block 3는 이미 두번째 부호화 사이클에서 *EOB*(End of Block)에 도달하였기 때문에 중요스캔에 포함이 되지 않고 Block 2의 AC2만이 부호화 된다. 이처럼 사이클 블록 부호화에서는 특정 변환블록 정보 전체를 부호화 후에 다음에 위치하는 변환블록들을 차례대로 부호화하지 않고, 화면 전체의 각 변환블록마다 동일한 부호화 사이클에 속하는 계수들을 동일한 부호화 및 전송 단위로 묶어서 순차적으로 부호화한다. 이렇게 할 경우 전송 대역폭이 갑자기 감소되었을 때, 주어진 비트량을 한 화면에 속하는 모든 변환블록에 고르게 할당할 수 있어 화면 전체의 화질을 고르게 유지할 수 있다. 사이클 블록 부호화에 대한 보다 자세한 설명은 참조문헌 [7], [8]에서 얻을 수 있다.

그림 4의 $\hat{E}^n$에 대하여 적용되는 부호화 사이클의 총 회수를 $K$로 가정하면, 사이클 인덱스(index)의 집합 $I$를 $I=\{0,1,2,...,K-1\}$로 표시할

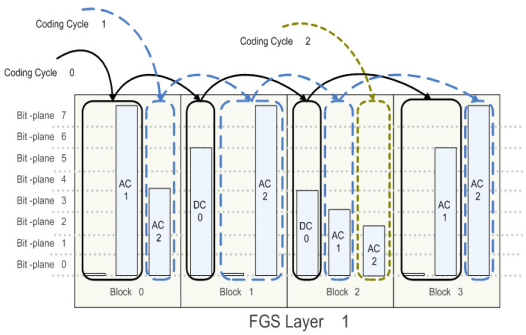



그림 5. SVC의 FGS 계층 부호화를 위한 사이클 블록 부호화 적용 예.



www.dbpia.co.kr

수 있다. 4x4 변환블록에서 $i$번째 부호화 사이클인 사이클 인덱스 $i$에 속하는 주파수 성분 $(u,v)$ 를 $(u_i,v_i)$로 표시한다면, 사이클 블록 부호화에 입력되는 신호 $\hat{E}^n$를 다음과 같이 표현할 수 있다.

$$\hat{E}^n(u,v)=\bigcup_{i\in I}\hat{E}^n(u_i,v_i) \tag{11}$$

한편, 사이클 블록 부호화에서 초기 사이클 인덱스에 포함되는 변환계수는 그렇지 못한 변환계수보다 일반적으로 화질 향상에 기여하는 중요도가 더 높다. 그리고, 전송 대역폭의 감소로 인해 FGS 계층에서 비트스트림 절삭(truncation)이 발생할 경우 부호화 정보가 복호기로 전송되는 순서가 상대적으로 앞서기 때문에, 복호기로 전달되는 확률 또한 상대적으로 높게 된다. 따라서, 초기 사이클 인덱스에 포함되는 변환계수에는 상대적으로 큰 값의 가중치 $\lambda$ $(0\le\lambda\le 1)$를 적용하여 고화질의 향상계층 참조화면 생성에 기여하는 비중을 높일 필요가 있고, 그렇지 못한 변환계수에는 작은 값의 $\lambda$를 적용하여 향상계층 참조화면 생성에서 그 영향력을 줄일 필요가 있다. 각 사이클 인덱스에 속하는 변환계수에 곱해지는 $\lambda$값을 결정하고 크기가 조절된 변환계수로 재생하는 과정은 그림 4에서 'Cycle adaptive coefficient reconstruction' 블록이 수행하게 된다. 결과적으로 $\hat{E}^n$ 신호가 'Cycle adaptive coefficient reconstruction' 블록을 통과후 재생되는 신호인 $\hat{E}^{n'}$는 다음과 같이 얻어진다.

$$\hat{E}^{n'}(u,v)=\bigcup_{i\in I}\hat{E}^n(u_i,v_i)\cdot\lambda_i \tag{12}$$

이 식에서 $\lambda_i$는 $i$번째 사이클 인덱스에 속하는 변환계수에 곱해지는 가중치이며, 기존의AR-FGS 기법에서는 그 값이 항상 1로 고정되어 있는 경우에 해당한다. $\lambda_i$는 사이클 인덱스 $i$에 속하는 변환계수가 FGS 복호기에 전송될 확률 $P_i$와 직접적으로 관련이 있기 때문에, 다음의 설명처럼 $P_i$와의 관계 규명을 통해 그 값을 설정할 수 있다.

어떤 변환계수의 원래의 완전한 재생 값을 $T$라고 가정하면, $\lambda$에 의해 크기가 조절 (scaling) 된 값은 $\lambda T$로 표시된다. 제안된 AR-FGS 부호기 구조에서 이 변환계수가 전송확률 $P$로 FGS 복호기에 전송되었을 경우 발생하는 왜곡 $D_1$은 다음과 같이 정의 된다.

$$D_1=(T-\lambda T)\cdot P \tag{13}$$

반면에, 변환계수가 확률 $(1-P)$로 FGS 복호기에 전송이 되지 않았을 경우 발생하는 왜곡 $D_2$는 다음과 같이 정의된다.

$$D_2=(0-\lambda T)\cdot(1-P)=\lambda T(P-1) \tag{14}$$

$D_2$의 경우에는 FGS 부호기에서 향상계층 참조화면 생성에 사용된 변환계수 정보가 FGS 복호기에 전혀 전달이 되지 않았기 때문에 부호기와 복호기의 향상계층 움직임 예측에 사용되는 참조화면 불일치로 인하여 화면을 통한 오류의 전파가 발생한다. 따라서, $D_2$의 왜곡에 보다 큰 가중치를 부여할 필요가 있고, 이 가중치 $w$ $(w\ge 1)$를 고려한 전체 왜곡에 대한 비용함수 (cost function) $D$는 다음과 같이 표현될 수 있다.

$$\|D\|=\|D_1\|+w\|D_2\| \tag{15}$$

이 비용함수를 최소화 하기 위한 방법은 다음 식과 같다.

$$\min(\|D\|)=\min(\|D_1\|+w\|D_2\|) \tag{16}$$

식 (16)을 만족하는 최적의 $\lambda$ 값은 최소평균제곱오차 (minimum mean square error)의 원리에 의해 다음의 방정식을 만족하는 해를 구함으로써 얻어낼 수 있다.

$$D_1\cdot P+w\cdot D_2\cdot(1-P)=0 \tag{17}$$

식 (17)을 풀게 되면, $\lambda$는 다음과 같이 얻어진다.

$$\lambda=\frac{P^2}{(P^2+w(1-P)^2)} \tag{18}$$

식 (18)에서 $w=3$으로 설정한 후 $P$와 $\lambda$ 사이



www.dbpia.co.kr

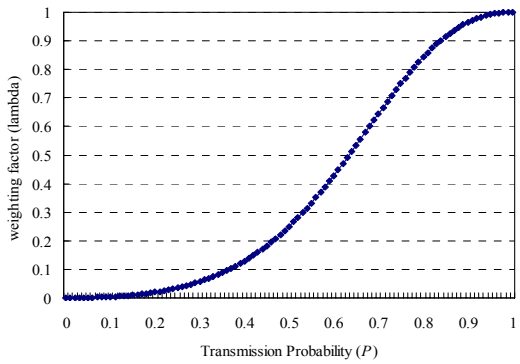



그림 6. 전송확률 $P$와 가중치 $\lambda$ 사이의 관계 ($w=3$인 경우).

의 관계를 그래프로 그려보면, 그림 6의 관계를 나타낸다. 그림 6에서 특수한 경우로서 $P=0$일 때는 $\lambda=0$이므로, 이 경우에는 향상계층의 참조화면 생성에 변환계수가 전혀 기여하지 못하도록 하고 있으며, 반면에 $P=1$일 때는 $\lambda=1$이므로, 이 경우에는 원래의 완전한 재생 변환계수값 $T$를 이용하여 향상계층의 참조화면 생성에 기여하도록 한다.

이상의 결과로부터 사이클 블록 부호화에서 사이클 인덱스 $i$에 해당하는 변환계수에 대한 전송확률 $P_i$가 주어질 경우 식 (18)의 관계를 통해 $\lambda_i$를 얻어낼 수 있다. 또한, 두 사이클 인덱스 $i$ 와 $j$ 사이에는 항상 다음의 관계식이 성립함을 알 수 있다.

$$P_i \geq P_j, \quad \text{if} \quad i<j \quad \text{and} \quad i,j \in I. \tag{19}$$

한편, $P_i$는 FGS 계층 비트스트림을 전송 서비스하면서 일정 시간 간격마다 그 값을 갱신할 수도 있고, 서비스 실시 이전에 사전 모의실험을 통하여 그 값을 설정할 수도 있다.

표 1은 CIF 규격으로 초당 15프레임으로 부호화된 Football, Mobile, Bus 테스트 영상 100 프레임을 256 Kbps로 전송하였을 때의 각 사이클 인덱스에 해당하는 변환계수의 부호화 정보가 복호기로 전송될 확률 $P_i$를 계산한 것이다. 실험에서 기본계층 부호화를 위하여 사용된 양자화 파라미터는 42이다. 표에서 사이클 인덱스가 증가할수록 해당 사이클 인덱스에 속하는 변환계수의 복호기 전송 확률이 감소함을 알 수 있다. 확률 $P_i$가 클 경우에는 그림 6에 의해 큰 값의 $\lambda_i$ 가 결정이 되며, 반대의 경우에는 0에 가까운 $\lambda_i$ 가 결정이 되어 제안된 방법에 적용이 된다.

표 1. 각 부호화 사이클 인덱스에 속하는 변환계수의 평균 전송 확률 ($P_i$)

| 부호화 사이클 인덱스 ($i$) | Football | Mobile | Bus |
|---|---|---|---|
| 0 | 1.0 | 1.0 | 1.0 |
| 1 | 0.984 | 0.996 | 0.973 |
| 2 | 0.912 | 0.931 | 0.938 |
| 3 | 0.854 | 0.876 | 0.832 |
| 4 | 0.635 | 0.721 | 0.713 |
| 5 | 0.521 | 0.618 | 0.578 |
| 6 | 0.438 | 0.504 | 0.487 |

## Ⅳ. 실험 결과

제안된 방법에 의한 AR-FGS 기법의 부호화 성능 향상을 검증하기 위하여 몇 가지 실험을 통한 결과를 제시한다. 실험에서는 그림 1의 부호화 구조를 따르는 JSVM 6.0(Joint Scalable Video Model 6) [6]의 표준 FGS 기법, 그림 2의 구조를 따르는 표준 AR-FGS 기법과 제안된 개선형 AR-FGS 기법의 부호화 성능을 비교한다. 실험에서는 CIF(352x288) 규격에 초당 15 프레임으로 부호화된 Football, Mobile, Bus 테스트 영상을 사용하며, 이들의 화면 구조는 B-화면은 포함하지 않으며, 첫 번째 화면이 I-화면으로 부호화되고 나머지 모든 화면은 P-화면으로 부호화 된다. 각 실험에서 기본계층 부호화를 위하여 사용된 양자화 파라미터 QP는 42또는 30이며, 그림 1과 그림 2처럼 1개의 FGS 계층이 기본계층을 기반으로 부호화 된다. 그리고, 표준 AR-FGS기법과 제안된 개선형AR-FGS 기법에서는 공통적으로 가중치 $\alpha=\frac{3}{16}, \beta=\frac{14}{16}$가 사용이 되었다.

그림 7은 세가지 부호화 기법으로 Football 영상을 부호화 하였을 때의 휘도성분에 대한 평균 PSNR 결과를 보여준다. 기본계층 부호화를 위하여 사용된 QP는 42이다. 그림에 보이듯이, 표준 AR-FGS 기법과 비교하였을 때, 제안된 AR-FGS 기법은 PSNR 성능을 최대 1.1 dB까지 향상시킬 수 있다. 특히, 중간 비트율 영역(mid-range)에서 최대의 PSNR 향상을 얻을 수 있다. 그러나, 비트율이 상위 비트율 영역으로 증가하게 되면, 표준 AR-FGS 기법에서 큰 값의 사이클 인덱스에 속하는 변환계수의 전송확률 또한 높아지므로, 표준



www.dbpia.co.kr

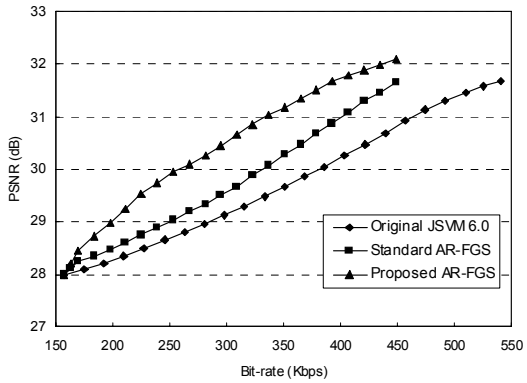



그림 7. Football 영상에 대한 JSVM 6.0의 FGS 기법, 표준 AR-FGS 기법, 그리고 제안된 AR-FGS 기법 간의 부호화 성능 비교.

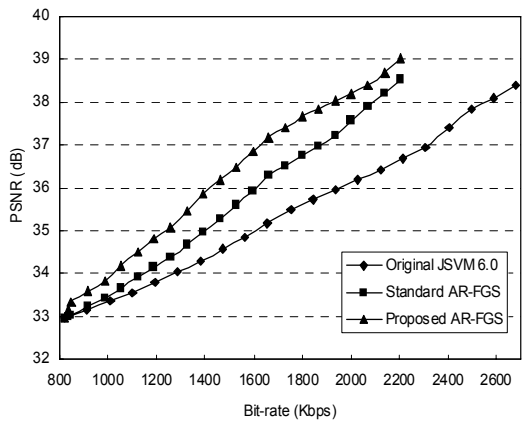



그림 8. Mobile 영상에 대한 JSVM 6.0의 FGS 기법, 표준 AR-FGS 기법, 그리고 제안된 AR-FGS 기법 간의 부호화 성능 비교.

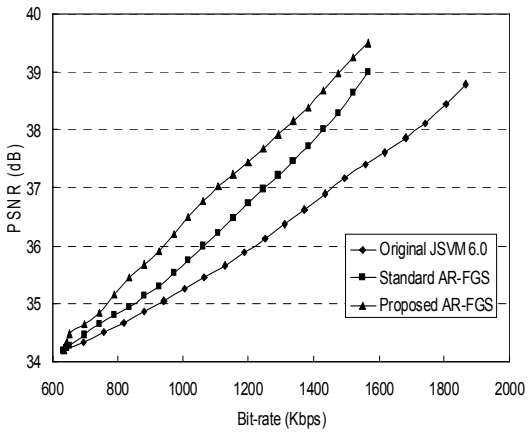



그림 9. Bus 영상에 대한 JSVM 6.0의 FGS 기법, 표준 AR-FGS 기법, 그리고 제안된 AR-FGS 기법 간의 부호화 성능 비교.

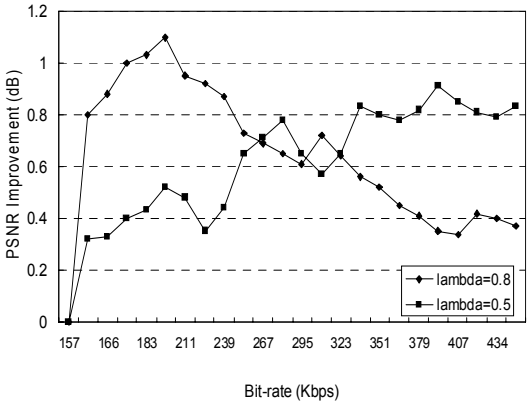



그림 10. $\lambda_i$=0.8, 0.5로 고정시킨 경우와 비교했을 때 제안된 방법의 PSNR 개선효과.

AR-FGS 기법의 PSNR 성능이 증대됨으로써 제안된 기법과의 부호화 성능 차이가 줄어들게 된다. 현재 상용 서비스 중인 지상파 DMB의 비디오 요구사항(requirement)이 CIF 규격에 512 Kbps 이하임을 고려할 때 200~400 Kbps 대역에서의 PSNR 성능 향상은 제안된 방법의 적용으로 DMB의 비디오 서비스 품질 향상을 기대할 수 있다.

그림 8과 그림 9는 각각 Mobile과 Bus 테스트 영상에 대해 그림 7과 동일한 목적의 실험을 수행하였을 때의 부호화 성능을 비교한다. 이 두가지 실험에서 기본계층 부호화를 위하여 사용된 양자화 파라미터 QP는 30으로 설정하였다. 그림 8에서는 표준 AR-FGS 기법 대비 최대 0.9 dB의 PSNR 이득을 얻을 수 있고, 그림 9에서는 최대 0.8 dB의 PSNR 이득을 얻을 수 있었다. 한편, 상위 비트율 영역으로 갈수록 제안된 방법에 의한 PSNR 이득은 점차 줄어들게 된다.

그림 10은 Football 영상에 대하여 $\lambda_i$값을 0.8과 0.5 로 고정시킨 경우와 비교했을 때, 전송 비트율 변화에 따른 제안된 방법의 PSNR 개선 효과를 비교한다. 그림에서 $\lambda_i$=0.8로 고정시킨 경우와 비교했을 때 낮은 비트율에서 PSNR 개선 효과가 두드러진다. 그 이유는 낮은 비트율에서는 큰 값의 사이클 인덱스에 속하는 변환계수들은 실제적으로 전송이 되지 못할 확률이 크기 때문에 제안된 방법에서는 $\lambda_i$에 작은 값이 할당되지만, $\lambda_i$=0.8로 고정시킨 경우에는 각 변환계수들이 $\lambda_i$의 큰 가중치로 인하여 향상계층 참조화면 생성에 크게 기여하기 때문이다. 반면에 $\lambda_i$ 값을 0.5로 고정시킨 경우에는 비교적



www.dbpia.co.kr

높은 비트율 영역에서 제안된 방법의 PSNR 개선이 두드러지게 나타나는 양상을 보인다.

## Ⅴ. 결 론

본 논문에서는 SVC 의 SNR 확장성에서 채택하고 있는 AR-FGS 기법의 부호화 성능 향상을 위한 개선된AR-FGS 기법을 제안하였다. 제안된 개선형 AR-FGS 기법에서는 FGS 부호화에 사용이 되는 사이클 블록 부호화의 기본 원리를 바탕으로, 사이클 블록 부호화에서 각 사이클 별로 생성되는 비트스트림이 향상계층 참조화면의 생성에 기여하는 중요도에 따라 가중치를 다르게 조절함으로써 특정 부호화 사이클에서 생성된 비트스트림 정보가 절삭에 의해 FGS 복호기에 전달되지 못 하더라도 복호화 시 그 영향을 최소화하여 화질 저하를 줄이는 방법을제안하였다. 제안된 기법의 적용으로 표준 AR-FGS 기법에 비하여 CIF 규격의 경우 1dB 안팎의 화질 개선 효과를 얻을 수 있음을 실험을 통해 확인하였다. 특히, 현재 상용 서비스 중인 지상파 DMB의 비디오 요구사항(requirement)이 CIF 규격에 512 Kbps 이하의 전송률임을 고려할 때 200~400 Kbps 대역에서 제안된 방법에 의한 1dB 안팎의 PSNR 성능 향상은 향후 SVC 기반의 DMB 서비스의 비디오 품질 향상에 크게 기여할 수 있다. DMB 서비스와 3세대 및 4세대 이동통신망의 출현으로 다양한 대역폭을 갖는 채널 환경에서 비디오 콘텐츠 전송 서비스를 실시할 필요성이 증대되고 있는데, 이러한 필요성을 가장 효과적으로 충족 시킬 수 있는 부호화 방식으로 SVC SNR 확장성 기술이 주목을 받고 있고, 제안된 방법을 SVC의 SNR 확장성에 적용할 경우 보다 고화질의 비디오 서비스 제공이 가능할 것으로 예상된다.

본 논문에서 제안된 개선형 AR-FGS 기법의 적용으로 FGS 의 고질적인 문제점인 부호화 효율 저하 문제를 효과적으로 해결할 수 있는 기반 기술 확보의 기틀을 확립하게 되었고, 향후 추가적인 부호화 성능 개선 연구에 본 논문의 결과물이 유용하게 활용이 될 수 있다.

## 참 고 문 헌


〔1〕 W. Li, "Fine granularity scalability in MPEG-4 for streaming video," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 1, pp. 299-302, Switzerland, May 2000.

〔2〕 M. Ghanbari, "Two-layer coding of video signals for VBR networks," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 771-781, June 1989.

〔3〕 W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 301-317, Mar. 2001.

〔4〕 J. Reichel, H. Schwarz, and M. Wien, "Scalable video coding- Working Draft 3," JVT-P201, July 2005, Poznan, Poland.

〔5〕 Yiliang Bao, M. Karczewicz, J. Ridge, X. Wang, "Improvements to fine granularity scalability for low-delay applications," JVT-O054, Apr. 2005, Busan, Korea.

〔6〕 J. Reichel, H. Schwarz, and M. Wien, "Joint Scalable Video Model 6," JVT-S202, Apr. 2006, Geneva, Switzerland.

〔7〕 J. Ridge, Y. Bao, M. Karczewicz, X. Wang, "Fine-grained scalability for H.264/AVC," in *Proc. IEEE Int. Symp. Signal Processing and Its Applications*, vol. 1, pp. 247-250, Aug. 2005.

〔8〕 J. Ridge, Y. Bao, M. Karczewicz, X. Wang, "Cyclical block coding," ISO/IEC JTC1/SC29/WG11, M11509, Jan. 2005.

〔9〕 W. Li, "Fine granularity scalability using bit-plane coding of DCT coefficients," ISO/IEC JTC1/SC29/WG11, M4204, Dec. 1998.

〔10〕 J. Macnicol, M. Frater, and J. Arnold, "Results on fine granularity scalability," ISO/IEC JTC1/SC29/WG11, M5122, Oct. 1999.

〔11〕 S. Li, F. Wu, and Y. Zhang, "Experimental results with progressive FGS coding," ISO/IEC JTC1/SC29/WG11, M5742, Mar. 2000.

〔12〕 S. Li, F. Wu, and Y. Zhang, "Study of a new approach to improve FGS video coding efficiency," ISO/IEC JTC1/SC29/WG11, M6779, Jan. 2001.

〔13〕 H. Huang, C. Wang, and T. Chiang, " A robust fine granularity scalability using trellis-based predictive leak," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 372-385, June 2002.

〔14〕 F. Wu, S. Li, and Y. Zhang, "A framework for efficient progressive fine granularity scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 332-344, Mar. 2001.

〔15〕 ISO/IEC JTC1/SC29/WG11, 14496-10 AVC, "Advanced video coding for generic audiovisual services," Mar. 2005.




www.dbpia.co.kr

**서 광 덕** (Kwang-deok Seo) 종신회원

1996년 2월 KAIST 전기및전자공학과 학사
1998년 2월 KAIST 전기및전자공학과 석사
2002년 8월 KAIST 전자전산학과 박사
2002년 8월~2005년 2월 LG전자 선임연구원
2005년 3월~현재 연세대학교 컴퓨터정보통신공학부 조교수
<관심분야> 영상부호화, 영상통신, 멀티미디어 통신시스템

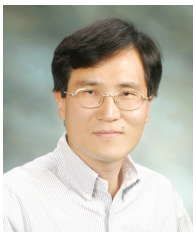
**정 순 흥** (Soon-heung Jung) 정회원

2001년 2월 부산대학교 전자공학과 학사
2003년 2월 KAIST 전자전산학과 석사
2005년 4월~현재 ETRI 방송미디어연구그룹 연구원
<관심분야> 영상부호화, 멀티미디어방송, 영상통신

**김 진 수** (Jin-soo Kim) 정회원

1991년 2월 경북대학교 전자공학과 학사
1993년 2월 KAIST 전기및전자공학과 석사
1998년 2월 KAIST 전기및전자공학과 박사
2000년 4월~현재 한밭대학교 정보통신컴퓨터공학부 부교수
<관심분야> 영상통신, 영상처리, 영상부호화

**김 재 곤** (Jae-gon Kim) 정회원



1990년 2월 경북대학교 전자공학과 학사
1992년 2월 KAIST 전기및전자공학과 석사
2005년 2월 KAIST 전자전산학과 박사
1992년~현재 ETRI 방송미디어 연구그룹 팀장
<관심분야> 비디오 신호처리, 영상통신, 디지털 방송, MPEG-7/MPEG-21



www.dbpia.co.kr