

영상 화면에 적용한 확장적응성 선택 방법에 관한 연구

송대건*

Research about Scalability selection method that apply to Image Scene

Song Dae Geon*

요 약

최근 네트워크대역 및 수신단말기의 해상도의 다양화에 의해 One Source-Multi use의 영상 전송을 구현하기 위한 기술로서 영상 확장적응성 기능이 주목받고 있다.

현재 MPEG에서는 H.264/MPEG-4 AVC를 기본으로 하는 SVC(Scalable Video Coding)의 표준화가 진행되고 있으며 SVC에서는 공간·시간·SNR 확장적응성을 제공할 수가 있다.

그러나 일반적으로 확장적응한 부호화는 비 확장적응한 부호화에 비해 부호화 효율이 떨어진다고 알려져 있다.

그래서 본 논문에서는 시간, SNR 확장적응성에 착안하여 SVC에 의한 확장적응성의 입자크기의 정도와 부호화 효율 관계를 기초로 하는 SNR 스케일러빌리티를 제공하는 FGS(Fine Granular Scalability) 기능을 전환함으로써 주관 품질의 향상을 꾀하는 수법을 검토한다.

ABSTRACT

An image scalability function is noted as technology to embody One Source-Multi-use's Image division by the network Bandwidth and diversification of resolution of reception terminal recently.

In MPEG present H.264/MPEG-4 AVC that do to standardization of SVC (Scalable Video Coding) that know go and SVC can offer space-time-SNR scalability.

But, encoding that do scalable usually is known that encoding efficiency drops than encoding that do rain scalable during time and treatise that see .

Examine technique to plan improvement of management quality as that aim to time, SNR scalability in treatise that see so and change FGS (Fine Granular Scalability) function that offer SNR scalability that do about scalability's particle size and encoding efficiency relation by SVC to foundation.

Keywords : MPEG-4 AVC, SVC, FGS, MCTF, IDCT,

I. 서 론

최근 네트워크대역 및 수신단말기의 해상도의 다양화에 의해 One Source-Multi use의 영상 전송을 구현하기 위한 기술로서 영상 스케일러빌리티 기능이 주목받고 있다.

현재 MPEG에서는 H.264/MPEG-4 AVC를 기본으로 하는 SVC(Scalable Video Coding)의 표준화가 진행되고 있으며[1] SVC에서는 공간·시간·SNR 스케일러빌리티를 제공할 수가 있다.

그러나 일반적으로 Scalable한 부호화는 비 Scalable한 부호화에 비해 부호화 효율이 떨어진다고 알려져있다[2]

그래서 본 논문에서는 시간, SNR 스케일러빌리티에 착안하여 SVC에 의한 스케일러빌리티의 입자크기의 정도와 부호화 효율 관계를 기초로 하는 SNR 스케일러빌리티를 제공하는 FGS(Fine Granular Scalability) 기능을 전환함으로써 주관 품질의 향상을 꾀하는 수법을 검토한다.

II. Scalable 동화상 전송을 위한 부호화 방법

1. Scalable 동화상 전송을 실현하기 위한 부호화 방법

* 경원대학교 IT대학 전자정보통신공학부 교수(dgsong@kyungwon.ac.kr)

Scalable 영상부호화 방식[4]을 Scalable 동화상 전송에 적용하였을 때 문제를 지적하여 새로운 부호화 방식에 대한 기본요건을 정리하고자 한다. 방식은 비계층 영상 정보로부터 기본 영상 품질을 할당하는 기본 계층과 기본 품질을 향상시키는 확장 계층을 생성하여 전송함으로써 스케일러빌리티를 실현시키는 것이다.

한편 전송효율에 착안하여 각 계층 간에 의존 관계를 갖기 위해서는 하위계층으로부터 순서적으로 우선제어할 필요가 있다.

이와 같은 경우 때문에 전송효율이 저하될 수도 있음을 고려해야 한다

1.1 Scalable 동화상을 전송하기 위한 시스템 구성

그림 1에 Scalable 동화상 전송을 위한 전송시스템 구성을 나타내었다.

본 시스템은 영상을 송출하는 서버와 전송 제어를 하는 라우터, 영상을 수신하는 리시버를 기본으로 구성되어있다.

카메라 등에서 촬영된 영상이나 미리 저장된 영상은 서버에서 전송 가능한 데이터로 변환되어 전송된다. 전송된 데이터는 라우터로부터 네트워크 상태에 따라 전송 제어가 실행되고 리시버에서는 이용환경에 따라 데이터를 수신하여 복호화함으로써 영상을 재생한다. 본 시스템으로부터 동일한 영상을 소스로 부터 다수의 영상품질을 제공할 수 있으며 보다 높은 스케일러빌리티를 실현할 수가 있다.

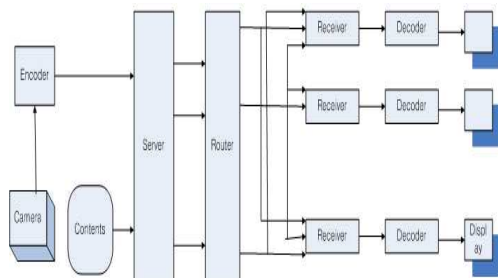


그림 1. Scalable 동영상 시스템
Fig 1. Scalable Moving Image System

1.2 Scalable 영상 부호화 방식의 특성

기존의 Scalable 영상 부호화 방식으로서 SNR 스케일러빌리티, MPEG-4FGS, 비트스트림 변환에 의한 계층 부호변환방식이 있다.

위의 방식을 품질 선택의 유연성, 부호화 효율, 부호기·복호기의 범용성, 의존성의 관점에서 정성적으로 비교하고자한다.

1.2.1 품질선택의 유연성

SNR 스케일러빌리티에서는 저 품질의 실현하는 기본품질과 품질 향상을 목적으로 하는 확장 계층으로 구성되어있으나 계층수가 유한적이기 때문에 품질선택에 있어서 제한이 있다.

또, 확장계층 데이터의 일부분이라도 부족하면 계층 전체에 에러가 확산되어 기본 화상 정도밖에 제공되지 않으며 따라서 보다 효율적인 품질을 선택하기 위해서는 확장 계층에 의한 스케일러빌리티를 실현할 필요가 있다. 이와 같은 요구를 만족시키기 위해서는 양자화수법의 다단화로부터 프레임마다 비트량을 조절할 수 있게 함으로서 확장계층에 대한 스케일러빌리티를 실현하여 일부 데이터가 부족하더라도 영향을 적게 하는 방식으로서 MPEG-4에 의한 FGS를 채택하고 있다.

1.2.2 부호기·복호기의 범용성

이 방식은 계층구조를 갖는 전용의 부호화기, 복호화기가 필요하다.

방법으로는 각 계층의 정보를 병렬로 복호화하고 역 양자화 하여 덧셈 연산한 후 IDCT를 해야 한다. 이 방식에서는 이미 비 계층구조에 부호화 끝난 영상정보에 대해서는 한번 복호화한 후 계층구조를 만들 필요가 있다. 이것에 비해 범용의 부호화기, 복호화기 사이에 있는 영상정보교환기(Transcoder)로부터 이미 부호화된 비계층 영상정보를 계층형 영상정보로 적은 처리량으로 변환할 수 있게 하는 방식이 검토되고 있다.

이 방식에서는 부호화기내에 계층구조를 생성할 뿐만 아니라 계층구조에서 변환기를 부호화기, 복호화기 사이에 두어 범용의 부호화기로서 이용가능하게 하고 있다.

1.2.3 부호화 효율

MPEG-2 SNR 스케일러빌리티에서는 기본 계층에 확장계층을 부가하였을 때 품질이 비계층 부호화 방식에 비해 저하됨이 보고되어있다.

또한 확장계층 정밀도를 부여한 MPEG-4 FGS에 있어서도 같은 특성을 갖고 있다. 이에 반해 비 계층방식과 같은 품질을 제공하

는 방식, 확장계층을 N층으로 확장한 방식은 확장 방식이 기본 계층과의 차이가 가역 복원할 수 있도록 설계되어 복원한 후에는 비계층 구조방식과 같은 부호량으로 같은 품질을 제공할 수 있게 한다.

또, 다계층 구조를 갖고 있는 방식에 있어서는 중간계층까지의 복원에 의해 제공된 품질은 필요로 하는 복원 부호량에 비해 저하되는 경향이 있다. 그래서 고위계층을 순서적으로 덧셈 연산한 경우에 한해서 선형적으로 영상품질을 향상시킬 검토가 필요하다.

1.2.4 의존성

기존의 방식으로부터 생성된 데이터는 전송 에러 등에 의해 중간 계층이 손실되었을 경우에 손실 계층 이상의 계층 데이터를 복호화 한다는 것은 불가능하다. 이런 경우를 대비하여 하위 계층의 우선순위를 높일 수 있는 복잡한 제어 구조가 필요하며 계층수의 증가에 따른 네트워크의 전송시 우선제어가 복잡하게 된다. 그림 2와 그림 3에서 예를 들어 5개 Layer에서 Layer2를 손실하였을 경우 Layer3, Layer4를 수신하였어도 최종적으로 생성된 품질은 Layer0, Layer1을 포함한 품질만 생성된다. 이와 같이 데이터 사이에 의존관계가 있는 경우에는 수신 부호량에 적합한 영상 품질을 얻기란 불가능하다.

그래서 데이터사이에 존재하는 관계를 배제하고 임의의 데이터를 조합함으로써 영상 재생을 할 수 있는 새로운 부호화 방식이 필요하리라 생각된다.

이로부터 손실된 데이터의 종류에 관계없이 수신부호량에 맞는 영상품질이 보정되기 때문에 데이터에 부합된 영상 품질의 제공이 가능하리라 생각된다.

이처럼 전송 데이터 구조를 서로 독립함으로써 전송함에 있어서 우선 제어를 간략화 하여 전송 에러에 대한 내성을 강화 할 수 있으리라 사료된다.

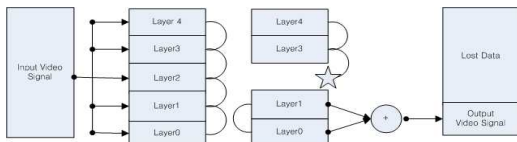


그림 2. 계층구조로의 변환과 복호
Fig 2 Conversion and decoder by class

structure

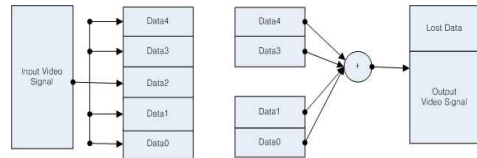


그림 3. 독립 데이터구조로의 변환과 복호
Fig 3. Conversion and decoding by independence data structure

III. 스케일러빌리티와 부호화 효율

1. 스케일러빌리티와 부호화 효율

SVC는 시간 스케일러빌리티를 실현하기 위하여 MCTF(Motion Compensated Filtering)을 이용한다.

MCTF는 시간방향의 서브 밴드 분할을 반복 실행함으로써 계층구조를 구성하고 시간 범용성을 실현하고 있다. MCTF에서는 GOP 사이즈에 따라서 획득 가능한 프레임율이 변화한다.

일반적으로 $2^N(N=0,1,2,3,...)$ 의 GOP 사이즈에서는 N계층의 시간 해상도가 얻어진다.

동화상의 시간적 복잡도나 동적 보상 예측 정도에 따라서는 Full decoder일대의 평균 PSNR의 최대치에 주어지는 GOP 사이즈가 다르다는 것을 이미 알고 있다.[3] 또, SVC는 SNR 스케일러빌리티를 실현하기 위하여 FGS를 사용하고 있다.

SVC는 FGS일 때에 Cyclic Block Coding이라 부르는 부호화로 실행하고 있으나 이것은 기본 Layer와 확장 Layer로 분할하기위하여 MPEG-4의 bit-plane Coding과 마찬가지로 비 FGS 부호화와 비교하여 약 1~1.5[dB]정도의 부호화 효율이 떨어짐을 알 수가 있었다. FGS에 의한 SNR 스케일러빌리티에서는 스테일러빌리티의 정밀도와 부호화효율에서는 서로 다른 관계가 존재하나 MCTF에 의한 시간 스케일러빌리티에서는 스케일러빌리티의 정밀도와 부호화 효율이 서로 다르다고는 단정 지을 수는 없다.

이와 같은 특성을 나타낸 것이 동화상의 시간 방향의 상관성에 의한 움직임에 대한 보상이 다르기 때문이다.

FGS를 이용한 경우 부호화 효율의 떨어짐은 주관품질의 떨어짐에 영향을 미치며 이때 주관품질의 떨어짐에 의한 영향력은 동화상의 특성에 따라 달라진다고 사료된다.

그래서 4절에서는 시간·SNR 스케일러빌리티를 교환함으로써 이와 같은 품질의 떨어짐을 피할 수 있는 방법에 대하여 검토하고자 한다.

2. 시간 공간 주관 품질 추정

주관평가실험에서 얻어진 평가자의 주관 평가의 평균치를 MOS(Mean Opinion Score)라 한다.

문헌 [4][5]에서는 각각의 공간 · 시간방향의 주관품질 추정식의 제안하였다

공간 방향의 주관품질은 다음식과 같이 표현할 수가 있다.

$$MOS_s(x) = at(x-t)^3 + bt(x-t) + c$$

(1)

$$t = b_1\epsilon + b_0$$

(2)

여기서, x 는 PSNR의 값, t 는 edge의 통계량을 나타내고 있다.

결국 식(1)은 edge의 통계량으로부터 공간방향의 주관적 품질량을 추정하는 식이다.

또, 시간방향의 주관적 품질은 다음과 같은 식으로 나타낼 수 있다.

$$f(m) = \frac{a-b \cdot (\log \frac{m}{c})^2 + 2.26}{3.58}$$

(3)

$$MOS_t = 6.0 - \alpha(f_{max})^\beta$$

(4)

여기서 $f(m)$ 는 평가함수라 부르며 프레임 삽입의 필요성을 나타내며 m 은 움직이는 벡터의 프레임내의 최대치를 나타내고 있다.

즉 식(4)는 움직이는 벡터의 프레임내 최대값으로부터 시간방향의 주관적 품질을 추정하는 식이다.

IV. 시간·SNR 스케일러빌리티 교환법

문헌[3]에 의하면 기존의 방법에서는 시간 ·

SNR 스케일러빌리티의 조합을 전제로 하고 있기 때문에 FGS 사용에 의한 분할손실이 고려되어 있지 않다.

그래서 제안한 시간 · SNR 스케일러빌리티의 교환 방법에서는 FGS ON/OFF 교환법을 함으로써 동화상의 특성에 따라 SNR · 시간 스케일러빌리티 중에서 어느 한쪽의 스케일러빌리티만을 갖고 있음으로 해서 스케일러빌리티를 유지해가면서 주관적 품질을 향상을 꾀한다.

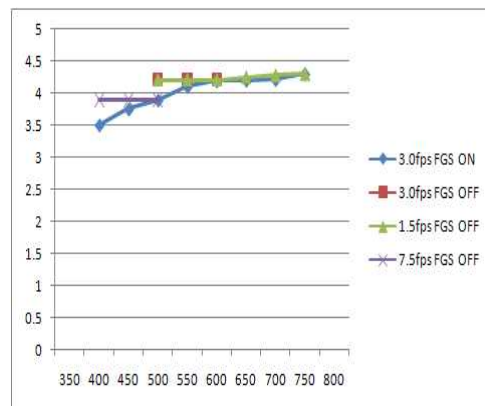
본 논문의 교환법에서는 스케일러빌리티를 다음과 같은 두 가지 중 하나만으로 구현하고자한다.

- 프레임율은 30[fps]로 고정하고 FGS를 이용함으로써 SNR방향의 스케일러빌리티만을 제공한다.

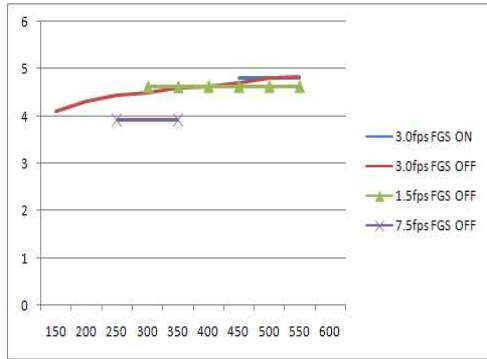
- FGS를 이용하지 않고 MCTF에 의해 구현한 시간 방향의 스케일러빌리티만을 제공한다.

우선, 이러한 2가지 스케일러빌리티로부터 구현된 동화상에 대하여 시간 · SNR 스케일러빌리티 교환법에 대하여 제 3절의 2가지 예측법을 이용하는 경우의 추정 MOS를 그림 4에 나타내었다.

그림 4에서는 비트율(bit rate)에 대한 연속적인 변화를 나타내고 있는 데이터 결과는 프레임율(frame rate)을 30[fps]로 고정하고 FGS를 이용하여 SNR 스케일러빌리티만으로 스케일러빌리티를 실현시킨 것이다.



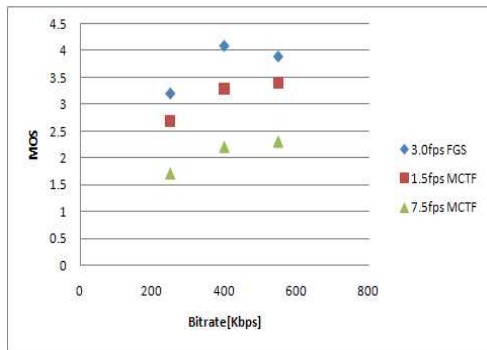
(a) FOOTBALL



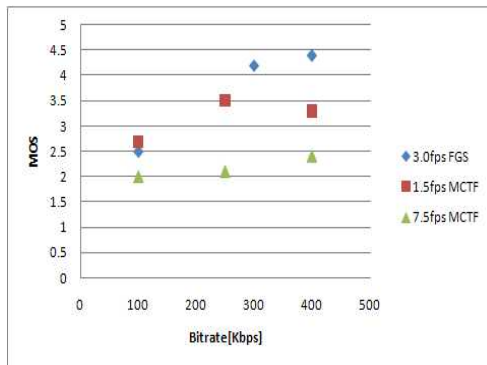
(b) HARBOUR

그림 4. 추정치에 의거한 추정 MOS

Fig 4. Conclusion MOS that conform to presume



(a) FOOTBALL



(b) HARBOUR

그림5. 주관평가실험에 의한 MOS

Fig 5. MOS by a Management Estimation Experiment

일반적으로 비트율에 대해서 이산적인 변화를 나타내고 있는 데이터 결과는 FGS기능을 이용하지 않고 MCTF를 이용한 시간 스케일러빌리티에

의한 스케일러빌리티를 실현한 것이며 위부터 아래 순으로 30[fps], 15[fps], 7.5[fps]로 되어있다.

그림 4(a)부터 추구와 같이 시간적 공간적으로 복잡한 큰 동화상에서는 프레임율이 떨어짐에도 불구하고 시간방향의 주관품질의 떨어짐이 비교적 적었으며 또한 프레임수를 적게 함에 따라서 1프레임당 소비하는 비트수가 증가하며 이러한 증가에 따라 PSNR의 상승값이 크게 때문에 7.5[fps]에 있어서도 시간 스케일러빌리티에 의해 실현하는 편이 높은 주관품질을 나타낼 수 있다고 생각된다.

한편, 그림 4(b)의 HARBOUR처럼 시간적 복잡도가 낮은 동화상에서는 프레임율이 떨어짐에 따라 시간 방향의 주관품질이 떨어짐이 현저하게 컸으며 또한 프레임수를 적게 함에 따라서 1프레임당 소비하는 비트량이 증가함에 따라서 PSNR 증가하나 그 상승 값은 작기 때문에 7.5[fps]에 의한 시간 스케일러빌리티 쪽이 주관품질이 낮은 결과가 나왔다고 생각된다.

또, 그림 5에 그림 4와 같은 조건의 동영상에 대하여 시간·SNR중 어느 하나만의 스케일러빌리티를 가진 경우의 주관평가실험 결과를 나타내었다.

그림5에 의하면 이번 주관 평가 실험 결과에서는 대부분의 동화상에 대하여 프레임율을 30[fps]로 고정시키고 FGS를 이용한 SNR 스케일러빌리티에 의한 스케일러빌리티를 함으로서 보다 높은 주관적 평가를 얻을 수가 있었다.

이것은 그림 4에서 나타난 결과와는 다른 결과 값을 나타내고 있다. 그 원인으로서는 다음 3가지로 생각해 볼 수가 있다.

첫째, 이번 실험의 주관적 평가 실험의 실시 방법에 있다고 생각한다.

이번 주관적 평가 실험에서는 이중자격법을 사용하고 있으나 이것은 프레임율 30[fps]의 원화상과 프레임율 7.5[fps], 15[fps], 30[fps]의 평가 대상 화상과서로 다른 평가의 결과를 얻었다는 점이며 프레임율을 적게 한 평가 대상 화상 쪽이 원래의 화상과의 차이가 현저하게 크게 나타난 원인이라고 생각 할 수 있다. 한편 평가 대상자에게 평가 대상 화상만을 보여주고 그 평가의 결과만으로서 구하는 단일자격법이 있다. 이 평가법에서는 평가자는 원 화상에 대하여 정보를 가지고 있지 않기 때문에 프레임율의 차이가 주관품질에 미치는 영향이 이중자격법에 의한 것보다는 다르게 평가대상화상에 대한 레이트를 보다 정확하게 얻을 수가 있지 않을까 생각된다.

두 번째는 식(4)에서의 시간 방향의 추정식의 이용 방법에 있다고 생각된다.

식(4)에서 주어져 있는 MOS_t 의 추정식은 간격이 다른 프레임율을 나열한 경우에 있어서의 주관적 품질을 예측하는 식으로서 이번 실험에서는 이 식을 같은 간격의 프레임율의 시간 방향의 주관적 품질의 예측에 이용하였다는 점이다.

세 번째로는 기존에 사용하고 있던 추정식의 특징량에 있다고 생각된다.

식(4)에서 특징량으로서 동적 벡터의 최대값을 사용하고 있다. 그러나 실제로는 물체가 이동하는 크기에 부가해서 이동하는 방향도 중요한 요소 중 하나라고 생각 해 볼 수 있다.

예를 들어 항구와 같이 등속 직선 운동을 하는 동화상에 대해서는 물체가 이동하는 방향이 일정하기 때문에 프레임율의 저하에 따른 주관적 품질의 영향이 그리 많지 않다고 생각된다.

반면에 축구와 같은 랜덤한 움직임을 가지는 동화상에서는 프레임율의 저하에 따른 주관적 품질의 영향이 많다고 생각된다.

V. 결 론

본 논문은 스케일러빌리티 동화상 전송 방식에서의 부호화 방식에 대하여 검토를 하였으며 시간 스케일러빌리티와 SNR 스케일러빌리티를 착안해 기존의 시간·SNR 스케일러빌리티 조합 방법에 대하여 이러한 스케일러빌리티를 교환함으로써 보다 효율적인 주관품질의 향상을 높이는 방법을 모색할수 목적으로 하였다.

먼저 스케일러블 영상부호화 방식을 적용하였을 경우의 방식은 영상 정보를 계층구조를 가진 데이터로 변환함으로써 스케일러빌리티 성질을 실현하는 방법이다. 그러나 서로간의 데이터가 존재하고 있기 때문에 전송 에러에 의해 발생한 중간 계층의 데이터가 손실되었을 경우 수신 부호량에 근접한 영상 품질을 얻어내기란 불가능하다. 이를 해결하기 위하여 전송할 때에 복잡한 구조의 우선 제어가 필요로 하며 계층수가 증가함에 따라 서버·네트워크에 대한 부하가 증대되어 전송 효율에 있어서도 저하를 동반하게된다.

그래서 이를 해결하기 위하여 부호화 방식에 대한 구할 수 있는 기본 조건들을 정리하여 데이터간에 존재관계를 배제하는 새로운 부호화 방식을 검토한 결과 앞의 조건들을 기본으로하여 비트스트림 변환에 의한 독립 데이터를 생성하는 부호화 방식에 대하여 검토 해야 할 필요가 있다고 생각된다.

또한 시간 스케일러빌리티와 SNR 스케일러빌리티의 관계를 착안해 기존의 시간·SNR 스케일러빌리티 조합 방법에 대하여서는 이런 스케일러빌리티들을 교환함으로써 보다 효율적인 주관 품질의 향상을 시키는 방법들이 모색되어야 된다고 생각되어 유효성을 확인하고자 각각의 스케일러빌리티만을 갖는 동화상을 만들어 주관품질 평가 실험을 한 결과 예상한 결과만큼은 도달되지는 못하였지만 평가의 결과 원인은 알 수가 있었다.

따라서 앞으로 이런 문제점들을 보완하기 위하여 단일 자격법등을 이용한 주관적 평가 시험 통해 결과를 토대로 새로운 특징량의 검토를 하고자 한다.

참고문헌

- [1] J.Reichel, H. Schwarz and M. Wien : "Scalable Video Coding-Working Draft 4", Joint Video Term(JVT)of ISO/IEC MPEG &ITU-TVCEG 16th Meeting, Doc, JVT-Q201, Oct. 2005
- [2] ISO/IEC 11172-2, International Standard : "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbitd/s", 1993
- [3] D.Morrison: "Video Transcoders with Low Delay", IEICE Trans.Comms., Vol E80-B No6 pp 963-pp969
- [4] G. Keerman, R.Hellinghuizen, F. Hoeksema and G Heideman : "Transcoding of MPEG Bitstreams", Signal Processing Image Communication Vol 8, 481-500, 1996
- [5] ISO /IEC 13818-1, IS "Info.Tech.- Generic Co ding of Moving Picture and Associted Audio Info.: System", 1995

저자약력



송대건(Song Dae Geon)

1981년 2월: 경희대학교 전자공학
학과 졸업
1988년 3월: 일본 와세다대학원
정보공학과 석사
1998년 2월 경희대학교 전자공
학과 박사수로
1989년 3월~현재 경원대학교 IT
대학 전자정보통신학부 전자공
학전공 교수

<관심분야> 정보보안, 신호처리, 영상처리,
임베디드 시스템 설계