

일반논문-08-13-1-17

## H.264 스케일러블 확장 규격의 고속 부호화를 위한 향상계층의 RD-최적화된 탐색 모드 제한 기법

이금희<sup>a)</sup>, 임진영<sup>a)</sup>, 정경훈<sup>a)</sup>, 김기두<sup>a)</sup>, 강동욱<sup>a)†</sup>

### RD-Optimized Search-Mode Restriction Method on the Enhancement Layer for Fast Encoding of H.264 Scalable Extension

Keum-Hee Lee<sup>a)</sup>, Jin-Young Lim<sup>a)</sup>, Kyeong-Hoon Jung<sup>a)</sup>, Ki-Doo Kim<sup>a)</sup>, and Dong-Wook Kang<sup>a)†</sup>

#### 요 약

본 논문에서는, 기본계층과 향상계층의 최적 모드 사이의 높은 상관성을 이용하여 기본계층의 모드에 기초하여 향상계층의 부호화 모드를 효율적으로 제한할 때, 생존하는 탐색 모드의 집합은 모드 제한에 의해서 발생할 추가 RD-비용을 최소화 하도록 결정되어야 함을 제시한다. 이 집합에 의거하여 H.264 스케일러블 확장 규격(H.264 SE)을 위한 고속 모드 결정 알고리즘을 구현하였다. 시뮬레이션 결과는 제안하는 알고리즘이 모드들의 조건부 확률에 기초한 기존의 모드 제한 기법에 비해서 RD-성능 면에서 우수함을 입증하고 있다.

#### Abstract

We investigate that when we restrict the encoding modes on the enhancement layer based on that of the base layer exploiting the high correlation between the best modes at base and enhancement layers, we should determine the set of surviving search modes to yield the minimum excessive RD-cost caused by the restriction. Based on these sets, we implement a fast mode decision algorithm for H.264 Scalable Extension (SE). Simulation results show that the proposed algorithm is superior in the sense of the RD-performance to the conventional mode restriction method which is based on the conditional probabilities of the best modes.

Keywords: H.264 SE, 고속 부호화, 모드 제한, 추가비용함수

#### 1. 서 론

최근 H.264|AVC의 확장 규격으로서 SVC(Scalable Video Coding)가 국제 표준으로 완성되었다<sup>[1]</sup>. 기존 표준과 비교할 때, SVC가 갖는 특징은 하나의 부호화된 비트스트림으

로부터 다양한 해상도와 프레임율로 복호 해낼 수 있도록 비트스트림을 계층적으로 부호화한다는 점이다. 이 특징을 SVC의 스케일러빌리티라고 한다. 스케일러빌리티를 갖는 SVM을 사용함으로써 동영상 콘텐츠에 대한 OSMU(one-source multi-use) 환경을 매우 효율적으로 구축할 수 있게 되었다<sup>[2]</sup>.

SVC의 스케일러빌리티는 입력 영상을 공간적으로는 계층으로 분해하고, 시간적으로는 계위적 B 픽처 구조를 활용하여 분해한 다음 계층별로 부호화함으로써 달성된다.

a) 국민대학교

Kookmin University

† 교신저자 : 강동욱(dwkwang@kookmin.ac.kr)

※ 이 연구는 2006년도 국민대학교 교내연구비 지원을 받아 수행된 연구임.

이때 공간적으로 가장 낮은 해상도의 계층을 기본계층(base layer)이라고 하며, 그 나머지 계층을 향상계층(enhancement layer)이라고 구분하며, 일반적으로 기본계층은 H.264/AVC와의 호환을 위하여 H.264/AVC 표준을 따라 부호화하고 있다.

SVC의 표준 코덱인 JSVM(Joint Scalable Video Model)은 부호화 과정에서 계층 사이의 상관성을 이용하기 위해서 기본계층(base layer)의 움직임 벡터와 기본계층의 복호화상을 향상계층(enhancement layer)의 예측 부호화에 이용하고 있다<sup>[1]</sup>. 즉, 기본계층의 움직임 벡터를 해상도 비율로 스케일링하여 향상계층 움직임 예측의 예측 벡터로 사용함으로써 향상계층 움직임 벡터 추정의 시간을 단축시키고 있으며, 기본계층의 복호화상을 보간하여 그 보간된 영상을 향상계층의 예측 영상으로 사용함으로써 예측 부호화의 효율을 높이고 있다.

부호화에 걸리는 시간의 증가는 스케일러블 비디오 부호의 가장 큰 단점이다. 부호화 효율을 높이기 위해 SVC에서는 RD(Rate Distortion) 비용함수를 정의하고, 각각의 매크로블록에 대해 RD 비용함수 값을 최소화하는 모드를 선택한다. 다시 말해, 모든 가능한 모드에 대해 RD 비용함수 값을 계산한 후 그 중에서 가장 작은 값을 갖는 모드를 해당 매크로블록의 부호화 모드로 결정하는 과정을 거치도록 하고 있다. 이와 같은 과정을 거치기 때문에 SVC는 높은 부호화 효율을 얻을 수 있지만 이로 인한 반복적인 연산과 복잡한 구조는 SVC의 실제적인 응용에 어려움을 주고 있다.

그러나 기본계층과 향상계층 영상 사이에는 매우 큰 상관성이 존재하기 때문에 그 부호화 모드 사이에도 또한 높은 상관성이 존재한다. 이 성질을 이용하면 기본계층에서 특정 모드가 선택되는 경우, 그 모드를 바탕으로 향상계층에서 탐색하는 모드를 적절히 제한함으로써 부호기의 전체적인 RD 성능을 크게 떨어뜨리지 않으면서도 부호화 시간을 크게 줄일 수 있다. 이를 실제로 구현한 부호화 속도 개선 알고리즘들이 최근 발표되고 있다<sup>[3][5]</sup>. He Li et al.는 화면간 모드 결정 시 하위 계층에 대한 모드를 판별하여 향상계층에서 후보 모드의 수를 줄여 나감으로써 고속 블록 모드 결정을 할 수 있음을 최초로 보였다. 이에 반하여

이범석 등은 기본계층의 움직임 벡터가 (0,0)일 때 향상계층에서는 무조건 INTER\_16x16으로 모드를 제한하여 RD-최적화를 통한 움직임 추정 및 모드 결정을 조기에 완료하는 알고리즘을 추가하여 He Li et al.의 알고리즘의 성능을 더욱 향상시켰다.

기존의 고속 부호화 알고리즘들은 일반적으로 다음과 같은 부호화 원리를 따르고 있다. 먼저 같은 프레임 번호를 갖는 기본계층 프레임이 존재하는 향상계층 프레임을 병치 프레임(collocated frame)이라고 정의하고, 그렇지 않은 프레임을 비병치 프레임(non-collocated frame)이라고 정의하자. 그러면 병치 프레임에 대해서는 대응하는 기본계층의 최적 매크로블록 모드에 대한 향상계층의 최적 매크로블록 모드의 조건부 발생확률(conditional probability)을 구하여 그 확률이 큰 몇 개의 모드만을 탐색함으로써 목표를 달성할 수 있다. 그리고 비병치 프레임에 대해서는 시간적으로 가장 인접하는 참조 프레임(향상계층)의 최적 매크로블록 모드에 대한 조건부 발생확률이나 그 참조 프레임들과 병치하는 (또는 시간적으로 일치하는) 기본계층 프레임의 최적 매크로블록 모드에 대한 조건부 발생확률을 사용할 수 있다. 한편 B-슬라이스의 경우에는 가장 인접한 2개의 참조 프레임과 시간적으로 일치하는 기본계층의 두 프레임, 즉 시간적으로 하나 앞선 프레임과 하나 뒤처진 프레임의 매크로블록 모드들에 대한 조건부 발생확률을 구하여 조건부 발생확률이 큰 몇 개의 모드만을 탐색함으로써 목표를 달성할 수 있다.

조건부 발생확률을 통한 향상계층에서의 매크로블록 모드 탐색 제한 기법은 직관적으로 이해하기 쉽고 설계하기에도 간편하다는 장점이 있지만 RD 성능의 입장에서 최적 솔루션을 제공하는 것은 아니다. 향상계층에서 매크로블록마다 탐색하는 모드의 수를 단 한 개로 제한하는 경우에는 조건부 발생확률이 가장 큰 하나의 모드를 선택하는 것이 가장 우수한 성능을 나타내리라는 것은 자명하다. 그러나 탐색하는 모드의 수를 늘려감에 따라 조건부 발생확률이 큰 순서대로 탐색 숫자만큼의 모드를 탐색하는 방법은 최적의 성능을 보장하지 못한다. 다음의 예를 통해서 그 이유를 알 수 있다. 예를 들어 조건부 발생확률이 큰 순서대로 모드를 나열할 때 A, B, C, D, E 모드의 순서이고, 그 확률

들이  $p_A \geq p_B \geq p_C > p_D \geq p_E$ 라고 가정하자. 이때 세 개의 모드만을 탐색하기로 한다면 조건부 발생확률 기법에 서는 모드 A, B, C만을 탐색하게 될 것이다. 그러나 모드들 사이에는 서로 보상적인 모드들도 존재한다. 예를 들어 모드 B가 최적 모드로 선택된 경우에 모드 B대신에 C로 부호화하여도 RD 비용함수 값이 크게 증가하지 않고, 그 반대로 모드 C가 최적 모드로 선택된 경우에 모드 C대신에 B로 부호화하여도 RD 비용함수 값이 크게 증가하지 않는 경우가 있다. 그런데 만일 모드 D가 최적으로 선택되는 경우는 비록 적지만 그 경우에 D 모드를 탐색하지 않고 모드 A나 B, C 중의 하나로 부호화할 때 RD 비용함수가 크게 늘어나는 경우가 많다면 어떻게 될까? 이 경우에는 모드 A, B, C 대신에 모드 A, B, D를 탐색하는 것이 결과적으로 더 우수한 RD 성능을 보이게 된다.

본 논문에서는 조건부 발생확률을 사용한 탐색 모드 제한 기법의 한계를 극복하는 새로운 탐색 모드 제한 기법을 제안하고, 이를 활용함으로써 RD 성능이 보다 개선된 새로운 고속 모드 결정 알고리즘을 제안한다. 다음 장에서는 모드 탐색 제한을 위한 새로운 기준을 제시하고, 이 기준에 따라 기본계층 모드별로 향상계층의 최적화된 탐색 모드 집합을 결정한다. 3장에서는 시뮬레이션을 통하여 이 모드 집합을 사용한 고속 모드 결정 알고리즘의 RD 성능 및 부호화 속도를 분석하고, 마지막으로 4장에서는 이에 따른 결

론을 짓는다.

## II. 추가비용함수의 정의와 이를 이용한 향상계층의 탐색 모드 제한 기법

### 1. 추가비용함수(Excessive Cost)의 정의

기본계층에 비해서 향상계층은 프레임율이 2배이며, 이에 따라 향상계층에는 그림 1과 같이 병치 프레임(collocated frame)과 비병치 프레임(non-collocated frame)이 번갈아 발생하는 경우를 가정하자. 또한 기본계층은 baseline profile을 갖는 H.264 부호기를 사용하여 부호화함으로써 I 또는 P-슬라이스로만 부호화되고 있다고 가정하자. 그리고 부호화 속도의 개선은 향상계층의 부호화 속도를 높임으로써 이루어진다고 가정하자. 그리고 향상계층의 부호화 속도를 높이기 위한 방법으로 향상계층에서는 매크로블록마다 가능한 모든 모드에 대해서 그 RD 비용을 계산하는 것이 아니라 제한된 모드들에 대해서만 해당 RD 비용을 계산하고, 그 중에서 최소의 비용을 보이는 모드를 해당 매크로블록의 부호화 모드로 결정하는 방식을 택한다고 가정하자. 이는 Li의 알고리즘을 포함하여 고속 부호화 방식에서 일반적으로 행하는 것을 가정하는 것으로서 매우 실용

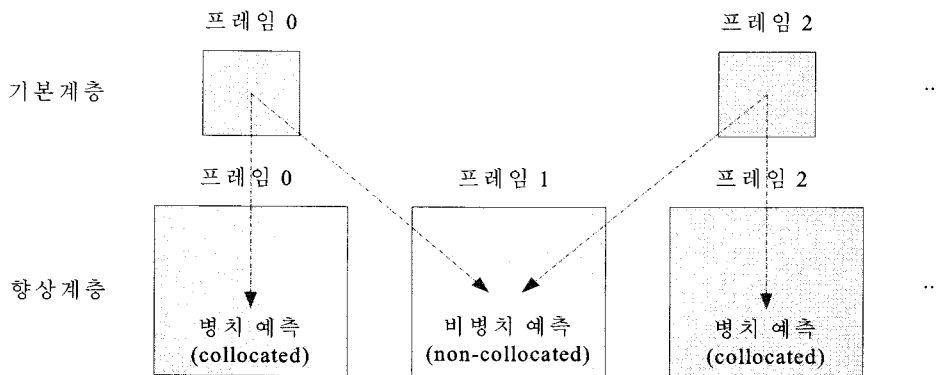


그림 1. 병치 프레임과 비병치 프레임

Fig. 1. Collocated frames and non-collocated frames in the enhancement layer

적인 가정이다.

항상계층의 부호화 모드 탐색을 제한하게 되면 부호화의 속도를 개선할 수 있는 반면에 RD 성능은 떨어지게 된다. 왜냐하면 최적의 RD 성능을 갖는 모드가 탐색 모드에 포함되지 못하는 경우에는 그 하위 성능을 갖는 모드가 선택될 것이고, 그것조차 탐색 집합에 없으면 차하위, 계속하여 차차하위 성능을 갖는 모드로서 부호화될 것이기 때문이다. 그러나 탐색할 모드를 제한하되 제한된 모드 집합에 최적의 RD 성능을 갖는 모드가 포함되어 있다면 그러한 문제는 발생하지 않을 것이다. 따라서 몇 개의 그리고 어떠한 모드들을 탐색 모드 집합에 포함시킬 것인가가 부호기의 부호화 속도와 RD 성능을 결정하게 된다.

먼저 부호화 모드  $m$ 에 대한 RD 비용함수  $C(m)$ 을 다음과 같이 정의하자.

$$C(m) = D(m) + \lambda R(m) \quad (1)$$

여기서  $R(m)$ 과  $D(m)$ 은 모드  $m$ 으로 부호화할 때 소요되는 비트수와 그 때 발생하는 복원영상에서의 왜곡(원영상과의 평균제곱오차)을 나타낸다.  $\lambda$ 는 라그랑지 승수(Lagrange Multiplier)이다. 이때 가능한 모드를 모두 탐색해보고 그중에서 최소 비용을 갖는 모드를 선택하는 경우에는 다음과 같이 최적의 최소 비용이 발생하게 된다.

$$C_{opt} = \min_m C(m) \text{ for all possible } m \quad (2)$$

시공간적으로 인접한 경우 기본계층과 항상계층의 매크로블록들은 최적 모드 사이에 큰 상관성을 보인다. 이 성질에 따라 기본계층의 모드에 따라 그 모드와 상관성이 큰 모드들로 항상계층의 탐색 모드를 제한하여도 RD 비용함수의 성능이 크게 떨어지지 않게 된다. 이제 이 성질을 이용하여 항상계층의 탐색 모드를 가장 제한하는 방법을 알아보자. 먼저, 기본계층의 해당 매크로블록에 대한 부호화 모드가  $b$ 일 때 항상계층에서 가능한 부호화 모드 중에서 임의로 선택한 집합  $S(b)$ 에 포함된 모드만을 탐색하는 경우 최종적인 비용이 다음과 같음을 이용하자.

$$C_{S(b)} = \min_m C(m) \text{ for all } m \in S(b) \quad (3)$$

이때 식 (3)으로 결정되는 비용과 식 (2)에서 주어진 최적 비용과의 차이는 모드 제한에 따른 추가되는 비용이므로 이를 다음과 같이 모드 집합  $S(b)$ 의 추가비용함수(Excessive Cost)라고 정의한다.

$$\Delta_{S(b)} = C_{S(b)} - C_{opt} \quad (4)$$

만일 탐색 모드를 제한하지 않고 모든 가능한 모드에 대한 RD 비용함수를 다 탐색해보고 부호화하는 경우에는 추가비용함수가 발생하지 않을 것이다. 즉,  $S(b) = \Omega$ (여기서  $\Omega$ 는 모든 모드의 집합)인 경우에는  $\Delta_{\Omega} = 0$ 이 된다. 반면에  $S(b)$ 에 포함된 모드의 수는 부호화 소요 시간과 직접적으로 비례한다. 따라서 부호화 속도를 높이려면 집합  $S(b)$ 의 크기를 최소화해야 하는데, 추가비용함수  $\Delta_{S(b)}$ 는  $S(b)$ 의 크기에 대해서 단조 감소하므로,  $\Delta_{S(b)}$ 를 매개로 하여 트레이드-오프 관계에 있는 항상계층 부호화의 RD 성능과 부호화 속도의 조절이 가능하다.

기본계층의 각 모드별로 항상계층의 탐색 모드 집합  $S(b)$ 들이 결정되어 있다면, 모든 모드를 다 조사해보는 최적의 경우와 비교할 때 주어진 모드 탐색 제한에 따른 항상계층 부호화의 전체적인 RD 비용 증가량은 다음과 같이 결정된다.

$$\Delta = \sum_b \Delta_{S(b)} \cdot p_b \quad (5)$$

여기서  $p_b$ 는 참조하는 기본계층 매크로블록의 부호화 모드가  $b$ 일 확률로서,  $\sum p_b = 1$ 이며 항상계층의 부호화되는 독립적으로 이미 결정되어져 있다고 할 수 있다.

그렇다면  $\Delta$ 를 정해진 목표 값 이하로 보장하기 위해서  $\Delta_{S(b)}$ 들을 어떻게 결정하는 것이 가장 유리할까?  $p_b$ 의 분포에 상관없이 목표 추가비용함수 값을 갖기 위한 간단한 방법은  $\Delta_{S(b)}$ 를 모두  $\Delta$  값 이하로 유지하는 것이다.

$$\Delta_{S(b)} \leq \Delta \text{ for all } b \quad (6)$$

다음 절에서는 식(6)을 이용하여 탐색 모드 집합  $S(b)$ 를 결정하는 과정을 예를 들어 소개한다.

### 2. 기본계층 참조 프레임 모드에 따른 향상계층 탐색 모드 제한

탐색 모드 제한 기법을 통해서 실현 가능한 부호화 속도는 이산적이다. 즉, 각 매크로블록마다 하나의 모드만으로 부호화할 때 최고의 속도를 보이고, 탐색하는 부호화 모드를 하나씩 늘려감에 따라 속도는 단조 감소한다. 참조하는 기본계층의 부호화 모드  $b$ 에 대해서 향상계층의 모든 가능한 모드의 수가  $M_b$ 라면  $M_b$ 가지의 부호화 속도로 부호화하는 것이 가능하다.

영상이 주어지면 이 영상을 실제로 부호화해봄으로써, 대응되는 기본계층 모드가  $b$ 인 모든 향상계층 매크로블록에 대해서 탐색 모드 집합의 크기를 1에서  $M_b$ 까지 증가시키면서 각각의 크기에서 최소의 추가비용함수 값  $\Delta_i(b), i = 1, 2, \dots, M_b$ 을 갖도록 하는 모드 집합  $S_i(b), i = 1, 2, \dots, M_b$ 를 결정할 수 있다. 즉,  $M = \max_b M_b$ 라고 할 때,  $i = 1, 2, \dots, M$ 에 대해서  $\max_b \Delta_i(b)$ 들을  $M$  개의 목표 추가비용함수로 정하여 식(6)과 결합하면, 각각의 모드  $b$ 마다 최대  $i$ 개의 모드만을 후보 모드로 포함하는 탐색 모드 집합  $S_i(b)$ 를 결정할 수 있게 된다.

표 1과 2은 CITY, FOOTBALL, MOBILE, BUS, HARBOUR, ICE, SOCCER, FOREMAN 등 8개의 영상을 부호화하면서 앞에서 설명한 과정을 통해서 얻은 탐색 모드 집합의 예를 나타내었다. 표 1에서 보이는 병치 향상계층에서는 시간적으로 일치하는 기본계층 P-슬라이스의 매크로블록의 모드를 조건으로 하여 조건부 추가비용함수와 탐색 모드 집합을 결정하고, 표 2에서 보이는 비병치 향상계층이란 시간적으로 하나 앞선 기본계층 P-슬라이스와 하나 뒤쳐진 기본계층 P-슬라이스(전자를 List 0, 후자를 List 1이라고 하였다)의 매크로블록 모드들을 조건으로 하여 조건부 추가비용함수와 탐색 모드 집합을 결정한다.

표 1과 2에서 첫 번째 열에 보인 표는 후보 모드를 최대 2개 허용한 경우이고 이때의 목표 추가비용함수 값은

2169.5 이다. 후보 모드를 2개 허용하였지만 기본계층의 모드에 따라 많은 경우 후보 모드를 1개만 사용하여도 그 조건부 추가비용함수 값이 목표 추가비용함수 값보다 작기 때문에 탐색 모드 집합의 크기는 1이 됨을 알 수 있다. 두 번째 열은 후보 모드를 최대 3개로 제한하는 경우에 해당되고, 세 번째는 4개로 제한한 결과이다. 이때의 목표 추가비용함수 값은 각각 585.9와 184.9이다.

표 1의 정보는 다음과 같이 의미를 파악할 수 있다. 두 번째 표에서 네 번째 행에 표시된 세 개의 동그라미는, 기본계층의 부호화 모드가 INTER\_8x16 모드일 때, 향상계층에서는 INTER\_16x16, INTER\_8x16, INTER\_8x8을 세 개의 후보 모드로 하여 탐색한다는 것을 의미한다. 반면에 표 2의 경우에는, 두 번째 행과 두 번째 열의 네 번째 행에는

표 1. 목표 비용함수와 기본계층 모드에 따른 병치 향상계층의 모드 탐색 범위  
Table 1. Modes to be searched in the collocated frames according to the target excessive cost and the corresponding base layer mode

		Target Excessive cost = 2169.5						
		EL candidate mode						
		16x16	16x8	8x16	8x8	I4	I16	I_BL
BL P- slice mode	Skip	○						
	16x16	○						
	16x8		○					
	8x16			○				
	8x8				○			
	I4	○						○
I16							○	○

		Target Excessive cost = 585.9						
		EL candidate mode						
		16x16	16x8	8x16	8x8	I4	I16	I_BL
BL P- slice mode	Skip	○						
	16x16	○		○				
	16x8	○	○	○				
	8x16	○		○	○			
	8x8	○		○	○			
	I4	○					○	○
I16	○						○	○

		Target Excessive cost = 184.9							
		EL candidate mode							
		16x16	16x8	8x16	8x8	I4	I16	I_BL	
BL P- slice mode	Skip	○	○	○					
	16x16	○	○	○					
	16x8	○	○	○	○				
	8x16	○	○	○	○				
	8x8	○	○	○	○				
	I4	○					○	○	○
I16	○						○	○	○

두 개의 동그라미가 표시되어 있는데, 그 의미는 다음과 같다. 두 번째 행으로부터 List 0의 모드가 INTER\_16x16이

고, 그 두 번째 열이므로 추가비용함수의 값이 목표 값인 585.9인 경우에 해당된다. 또한 네 번째 행은 List 1의 모드

표 2. 목표 비용함수와 기본계층 모드에 따른 비병치 향상계층의 모드 탐색 범위

Table 2. Modes to be searched in the non-located frames according to the target excessive cost and the corresponding base layer mode

Target Excessive cost = 2169.5		Target Excessive cost = 585.9		Target Excessive cost = 184.9		Target Excessive cost = 184.9	
List 0	EL candidate mode	List 0	EL candidate mode	List 0	EL candidate mode	List 0	EL candidate mode
Skip	16x16 16x8 8x16 8x8 I4 I16	Skip	16x16 16x8 8x16 8x8 I4 I16	Skip	16x16 16x8 8x16 8x8 I4 I16	16x16	16x16 16x8 8x16 8x8 I4 I16
List 1	Skip	○	○	Skip	○	Skip	○
	16x16	○	○	16x16	○	16x16	○
	16x8	○	○	16x8	○	16x8	○
	8x16	○	○	8x16	○	8x16	○
	8x8	○	○	8x8	○	8x8	○
	I4	○	○	I4	○	I4	○
I16	○	○	I16	○	I16	○	
List 1	16x16	○	○	Skip	○	Skip	○
	Skip	○	○	16x16	○	16x16	○
	16x16	○	○	16x8	○	16x8	○
	16x8	○	○	8x16	○	8x16	○
	8x16	○	○	8x8	○	8x8	○
	8x8	○	○	I4	○	I4	○
I4	○	○	I16	○	I16	○	
I16	○	○					
List 1	16x8	○	○	Skip	○	Skip	○
	Skip	○	○	16x16	○	16x16	○
	16x16	○	○	16x8	○	16x8	○
	16x8	○	○	8x16	○	8x16	○
	8x16	○	○	8x8	○	8x8	○
	8x8	○	○	I4	○	I4	○
I4	○	○	I16	○	I16	○	
I16	○	○					
List 1	8x16	○	○	Skip	○	Skip	○
	Skip	○	○	16x16	○	16x16	○
	16x16	○	○	16x8	○	16x8	○
	16x8	○	○	8x16	○	8x16	○
	8x16	○	○	8x8	○	8x8	○
	8x8	○	○	I4	○	I4	○
I4	○	○	I16	○	I16	○	
I16	○	○					
List 1	8x8	○	○	Skip	○	Skip	○
	Skip	○	○	16x16	○	16x16	○
	16x16	○	○	16x8	○	16x8	○
	16x8	○	○	8x16	○	8x16	○
	8x16	○	○	8x8	○	8x8	○
	8x8	○	○	I4	○	I4	○
I4	○	○	I16	○	I16	○	
I16	○	○					
List 1	Intra 4	○	○	Skip	○	Skip	○
	Skip	○	○	16x16	○	16x16	○
	16x16	○	○	16x8	○	16x8	○
	16x8	○	○	8x16	○	8x16	○
	8x16	○	○	8x8	○	8x8	○
	8x8	○	○	I4	○	I4	○
I4	○	○	I16	○	I16	○	
I16	○	○					
List 1	Intra 16	○	○	Skip	○	Skip	○
	Skip	○	○	16x16	○	16x16	○
	16x16	○	○	16x8	○	16x8	○
	16x8	○	○	8x16	○	8x16	○
	8x16	○	○	8x8	○	8x8	○
	8x8	○	○	I4	○	I4	○
I4	○	○	I16	○	I16	○	
I16	○	○					
List 1	Intra 16	○	○	Skip	○	Skip	○
	Skip	○	○	16x16	○	16x16	○
	16x16	○	○	16x8	○	16x8	○
	16x8	○	○	8x16	○	8x16	○
	8x16	○	○	8x8	○	8x8	○
	8x8	○	○	I4	○	I4	○
I4	○	○	I16	○	I16	○	
I16	○	○					

가 INTER\_8x16인 경우의 항상계층 탐색 모드들을 나타내게 된다. 그 행에 표시된 두 개의 동그라미는 이 경우에는 항상계층에서는 INTER\_16x16과 INTER\_8x16을 두 개의 후보 모드로 하여 탐색하여도 추가비용함수보다 보다 작거나 같은 왜곡만이 발생하리라는 것을 의미한다. 또 표 1에서 보이는 바와 같이 최대 허용 후보 모드의 수를 3으로 하여 목표 추가비용함수를 585.9로 정할 때, 병치 프레임에서는 기본계층의 모드가 MODE\_SKIP 이거나 INTER\_16x16일 때는 각각 후보 모드의 수를 1개와 2개로 더욱 제한하여도 된다. 또한 비병치 프레임에서는 List 0와 List 1의 모드 조합 49 가지 중에서 단지 22 경우에만 3개의 후보 모드를 탐색하고, 15 가지의 조합에 대해서는 단지 1개의 후보 모드를 나머지 12 가지의 조합에 대해서는 2개의 후보 모드만을 탐색하여도 됨을 보여주고 있다.

### III. SIMULATION RESULTS

JSVM 9.1<sup>[6]</sup>을 사용한 실험환경을 표 3에 나타내었다. 다양한 특성을 가지는 CITY, FOOTBALL, MOBILE, BUS, HARBOUR, ICE, SOCCER, FOREMAN 8 가지 영상을 탐색 모드 집합 결정을 위한 훈련과 RD 성능 분석에 모두 이용하였다.

표 3. 실험 환경  
Table 3. Conditions of experiments

GOP 크기	2
움직임 벡터 추정 범위	16 화소
움직임 벡터의 정밀도	1/4 화소
움직임 추정 함수	화소 벡터 추정 : SAD 부화소 추정 : SATD
공간적 스케일러블러티	2 계층
FGS 스케일러블러티	사용하지 않음
입력 영상	기본계층: QCIF 15fps 항상계층: CIF 30fps

표 4는 최대 탐색 허용 후보 모두 수에 따른 부호화 소요 시간의 감소를 보여준다. 항상계층의 가능한 매크로블록

모드는 모두 7 가지이므로 후보 모드의 수를 7에서 시작하여 1로 줄여감에 따라 부호화에 걸리는 시간은 차츰 줄어든다. 표 4에서 알 수 있는 바와 같이 MB(매크로블록)당 최대 탐색 후보 모드의 수를 줄여감에 따라 QP 값에 상관없이 비슷한 속도 개선 효과가 나타나기 시작한다. 탐색 모드 제한에 따른 평균적인 부호화 시간 감축 효과는 최소 9.2%에서 최대 65.4%이다.

표 5는 최대 탐색 허용 후보 모두 수에 따른 부호 비트수의 증가 정도를 보여준다. 항상계층의 후보 모드의 수를 7에서 시작하여 1로 줄여감에 따라 부호 비트수는 차츰 늘어난다. 탐색 모드 제한에 따른 평균적인 부호 비트수 증가는 최소 0.09%에서 최대 4.01%이다.

표 6은 최대 탐색 허용 후보 모두 수에 따른 양자화 잡음의 증가 정도를 보여준다. 항상계층의 후보 모드의 수를 7에서 시작하여 1로 줄여감에 따라 양자화 잡음은 차츰 늘어난다. 탐색 모드 제한에 따른 평균적인 양자화 잡음 증가는 최소 0.26%에서 최대 4.01%이다.

표 4. 탐색 모드의 수 제한에 따른 부호화 소요 시간 감축  
Table 4. Reduced encoding time according to the restriction of modes

(단위: %)

QP	MB당 최대 탐색 후보 모드의 수					
	6	5	4	3	2	1
36	12.8	30.8	45.5	57.9	65.4	65.8
32	11.3	28.0	44.6	56.4	65.7	66.2
28	6.8	23.1	41.8	53.6	64.3	65.2
24	5.9	20.9	41.3	52.5	64.2	64.6
평균	9.2	25.7	43.3	55.1	64.9	65.4

표 5. 탐색 모드의 수 제한에 따른 부호 비트수의 증가  
Table 5. Increased bit rate due to the restriction of modes

(단위: %)

QP	MB당 최대 탐색 후보 모드의 수					
	6	5	4	3	2	1
36	0.32	0.21	0.42	1.55	3.41	3.56
32	0.06	0.01	0.17	1.32	3.50	3.75
28	0.05	0.07	0.36	1.43	4.06	4.34
24	0.03	0.07	0.47	1.39	4.10	4.38
평균	0.12	0.09	0.35	1.42	3.77	4.01

표 6. 탐색 모드의 수 제한에 따른 양자화 잡음의 증가  
Table 6. Increased quantization noise due to the restriction of modes

QP	MB당 최대 탐색 후보 모드의 수 (단위: %)					
	6	5	4	3	2	1
36	0.75	0.36	1.47	2.45	4.25	4.42
32	0.37	0.25	1.06	1.86	3.79	4.01
28	0.11	0.22	0.92	1.74	3.53	3.72
24	0.21	0.20	1.08	1.77	3.68	3.87
평균	0.36	0.26	1.13	1.95	3.81	4.01

표 1과 2에서 보인 탐색 모드 집합은 훈련영상에 따라 달라지며, 그에 따라 부호화 성능도 달라진다. 표 7-9는 각각 FOOTBALL, FOREMAN, AKIYO 영상에 대해서 최적의 향상계층의 탐색 모드 집합을 결정하여 같은 영상들을 부호화할 때, 8개의 영상을 모두 훈련영상으로 하여 최적 탐색 모드 집합을 결정한 표 1과 2를 사용하여 부호화하는 것과 비교하여 RD 성능이 얼마나 개선되는지를 보여준다. FOOTBALL에 최적화한 알고리즘은 QP가 커질수록 부호 시간 단축 비율이 커진다. FOREMAN에 최적화한 알고리즘은 QP에 관계없이 부호화 시간이 9% 이상 단축되었고, AKIYO에 최적화한 알고리즘은 QP가 작을수록 부호화 시간이 최소 11.56%, 최고 18.83%까지 단축된다.

표 7. 8개의 훈련영상 대신에 FOOTBALL 영상에 최적화된 탐색 모드 집합을 사용하여 동일 FOOTBALL 영상을 부호화함에 따른 RD 성능 향상  
Table 7. RD-performance improvement by encoding FOOTBALL with the mode set optimized for FOOTBALL instead of overall 8 sequences

QP	부호 시간 단축 (%)	비트 레이트 감소 (%)	복원화질 향상 (dB)
40	13.75	1.93	0.01
30	10.01	3.54	-0.05
20	11.69	2.54	0.00
10	8.20	1.37	0.13

표 8. 8개의 훈련영상 대신에 FOREMAN 영상에 최적화된 탐색 모드 집합을 사용하여 동일 FOREMAN 영상을 부호화함에 따른 RD 성능 향상  
Table 8. RD-performance improvement by encoding FOREMAN with the mode set optimized for FOREMAN instead of overall 8 sequences

QP	부호 시간 단축 (%)	비트 레이트 감소 (%)	복원화질 향상 (dB)
40	10.12	-1.29	0.09
30	9.39	0.19	-0.01
20	9.93	0.17	-0.02
10	9.05	0.17	0.00

표 9. 8개의 훈련영상 대신에 AKIYO 영상에 최적화된 탐색 모드 집합을 사용하여 동일 AKIYO 영상을 부호화함에 따른 RD 성능 향상

Table 9. RD-performance improvement by encoding AKIYO with the mode set optimized for AKIYO instead of overall 8 sequences

QP	부호 시간 단축 (%)	비트 레이트 감소 (%)	복원화질 향상 (dB)
40	11.56	1.49	-0.04
30	13.71	2.77	-0.02
20	16.56	2.86	-0.03
10	18.83	0.59	0.08

표 10에서는 8개의 영상을 모두 훈련영상으로 사용하여 탐색 모드 집합을 결정하고, 그 집합을 사용하여 동일한 8개의 영상을 실제로 부호화하여 얻은 경우에 제안된 모드 제한 기법과 Li의 고속 모드 결정 기법의 성능을 비교하였다. JSVM 표준 부호화 알고리즘을 기준으로 두 알고리즘을 적용하는 경우의 부호 시간 단축과 그에 따른 부호 비트 수의 증가 비율 및 복원화질의 열화 정도를 서로 비교하였다. Li의 알고리즘은 MB당 2개에서 5개 사이의 모드를 후보 모드로 탐색하고 있어서 본 논문에서 제안하는 MB당 최대 4개까지의 모드를 탐색하는 경우와 비슷한 시간이 걸린다. 이 경우에 대해서 두 알고리즘에 따른 부호 비트수의 증가 비율과 양자화 잡음의 증가에 따른 복원화질의 저하 정도를 서로 비교한 것이다. 모든 부호기는 QP를 30으로 고정하여 부호화한 결과를 비교하였다. Li의 알고리즘에 비해 부호화 시간이 1.64초 정도 늘어났지만, JSVM과 비교하여 제안하는 알고리즘은 41.25%, Li의 알고리즘은 42.57%의 부호 시간 단축을 보이므로, 매우 유사한 부호화 속도로 부호화하고 있다고 할 수 있다. 비트 레이트는 JSVM과 비교하여 제안하는 알고리즘은 0.47%, Li의 알고리즘은 3.15% 증가하고, 복원영상의 화질은 각각 0.05dB와 0.11dB 열화됨을 보여준다. 따라서 제안하는 알고리즘을

표 10. Li et al.<sup>[9]</sup>의 알고리즘과의 부호 시간 및 RD 성능 비교  
Table 10. Comparison of encoding time and RD-performance of the proposed algorithm with those of Li et al.<sup>[9]</sup>

알고리즘	부호 시간(s)	비트 레이트(kbps)	복원화질(dB)
JSVM	124.56	810.70	33.08
Li의 고속 모드 결정 기법	71.53	863.09	32.97
제안하는 알고리즘	73.18	818.52	33.03



동일한 부호화 속도를 갖는 Li의 알고리즘과 비교하면, 비트 레이트는 5.16% 줄이면서 복원영상의 화질은 0.06 dB 향상시키는 결과를 보여준다. 이는 추가비용함수를 최소화하는 모드를 결정하는 것이 조건부 발생확률에 따라 모드를 결정하는 것보다 우수한 성능을 보일 것이라는 본 논문의 가설이 유효함을 증명하는 것이다.

#### IV. CONCLUSION

이 논문에서 우리는 기본계층과 향상계층의 시공간적으로 일치하거나 인접한 매크로블록들의 부호화 모드 사이에 높은 상관성이 존재한다는 성질을 이용하여, 기본계층에서 특정 모드가 선택되는 경우 그 모드를 바탕으로 향상계층에서 탐색하는 모드를 RD 성능의 면에서 최적으로 제한하는 방법을 제안하였다. 탐색 모드의 제한에 따른 RD-성능의 열화를 추가비용함수를 정의하여 대표하도록 하였으며, 추가비용함수를 통하여 서로 트레이드-오프의 관계에 있는 부호화 속도의 개선과 RD-성능의 저하 사이의 최적의 솔루션을 찾는 방법을 제시하였다. 8개의 서로 다른 특성을 갖

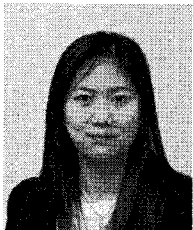
는 동영상에 대한 시뮬레이션을 통하여 제안하는 방법에 따라 탐색 모드를 제한하는 것이 부호화의 전체적인 RD 성능을 떨어뜨리지 않으면서도 부호화하는데 걸리는 시간을 기존 방법보다 더 효과적으로 줄일 수 있다는 것을 검증하여 보였다.

#### 참 고 문 헌

- [1] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), "Joint Draft 10 of SVC Amendment," JVT-W201, San Jose, Apr. 2007.
- [2] J.-R. Ohm, "Advances in scalable video coding," Proc. IEEE, vol. 93, no.1. pp. 42-56, Jan. 2005.
- [3] H. Li, Z. G. Li, and C. Wen, "Fast Mode Decision for Spatial Scalable Video Coding," Proc. ISCAS, pp.3005-3008, May 2006.
- [4] \_\_\_\_\_, "Fast Mode Decision Algorithm for Inter-Frame Coding in Fully Scalable Video Coding," IEEE Trans.on CASVT, vol.16, no.7, pp.889-895. Jul. 2007.
- [5] 이범식, 김문철, 함상진, 박창섭, 박근수, "공간적 스케일러블 비디오 부호화에서 계층간 모드 고속 결정 방법," 방송공학회논문지, 제12권 4호, pp. 360-372, 2007년 7월.
- [6] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), "JSVM 10 Software," JVT-W203, San Jose, Apr. 2007.

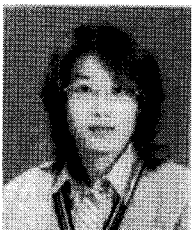
---

#### 저 자 소 개



##### 이 금 희

- 2006년 2월 : 국민대학교 전자공학부 졸업
- 2006년 3월 ~ 현재 : 국민대학교 전자공학부 석사 과정
- 주관심분야 : 영상압축, 영상신호처리



##### 임 진 영

- 2007년 2월 : 국민대학교 전자공학부 졸업
- 2007년 3월 ~ 현재 : 국민대학교 전자공학부 석사 과정
- 주관심분야 : 디지털영상처리, 멀티미디어신호처리

---

저 자 소 개

---



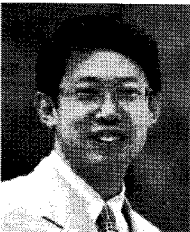
**정 경 훈**

- 1987년 2월 : 서울대학교 전자공학 졸업
- 1989년 2월 : 서울대학교 전자공학과 석사
- 1996년 2월 : 서울대학교 전자공학과 박사
- 1991년 12월 ~ 1997년 2월 : 한국영상산업진흥원 선임연구원
- 1997년 3월 ~ 2005년 2월 : 한동대학교 전산전자공학부 교수
- 2005년 3월 ~ 현재 : 국민대학교 전자공학부 교수
- 주관심분야 : 멀티미디어신호처리, 디지털 방송



**김 기 두**

- 1980년 2월 : 서강대학교 전자공학과 졸업
- 1988년 8월 : The Pennsylvania State University, MS(Electrical Eng.)
- 1990년 12월 : The Pennsylvania State University, MS(Electrical Eng.)
- 1980년 3월 ~ 1985년 12월 : 국방과학연구소 연구원
- 1998년 3월 ~ 1999년 2월 : 미국 UCSD, Visiting Scholar
- 1991년 3월 ~ 현재 : 국민대학교 전자공학부 교수
- 주관심분야 : 디지털통신, 디지털신호처리



**강 동 욱**

- 1986년 2월 : 서울대학교 전자공학 졸업
- 1988년 2월 : 서울대학교 전자공학과 석사
- 1995년 2월 : 서울대학교 전자공학과 박사
- 2000년 9월 ~ 2001년 8월 : Lucent Technology MTS
- 1995년 ~ 현재 : 국민대학교 전자공학부 교수
- 주관심분야 : 비디오 코딩, 영상통신