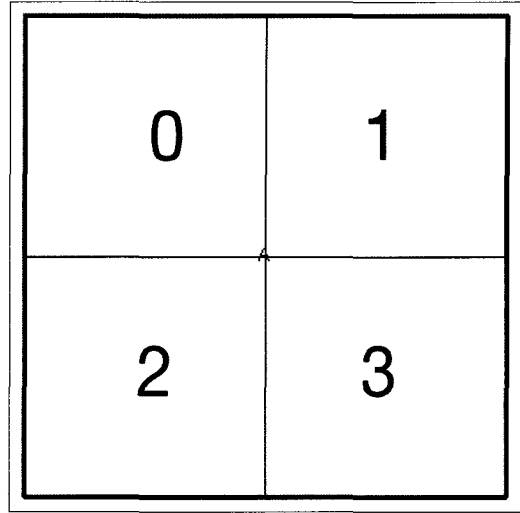




Draft)를 기준으로 하여, 현재까지 JVT에서 진행된 표준화 작업안 중에서 변환 및 양자화 과정에 대해서 자세히 설명하고자 한다[1].

H.264 | MPEG-4 Part 10에서는 H.263, MPEG-4등과는 달리 H.264 | MPEG-4 Part 10 표준은 정수 변환(integer transform)을 채택함으로써 인해서 변환 연산 수행시에 부족한 분해능(resolution)에 따라서 발생할 수도 있는 오차(mismatch)를 근원적으로 제거하였다. 즉, 기존 동영상 및 정지영상 표준에서 사용하였던 변환인 이산 여산 변환(DCT: Discrete Cosine Transform) 연산이 부동 소수점 연산(floating point operation)을 채택함으로써 인해서, 개개의 구현에 따라서 변환 연산의 결과가 달라질 수 있는 여지가 남겨져 있었다. 그러나 H.264 | MPEG-4 Part 10에서는 상기한 바와 같이 정수 연산 및 비트 이동(bit shift) 연산만으로 변환을 정의함으로써 인해서 디지털 시스템에서 구현시에 연산으로 인한 오차가 날 수 있는 여지를 표준화 작업 중에 근원적으로 제거하였다.

본 논문의 2장에서는 변환을 위해서 매크로블록에서 휘도 및 색상 신호의 구획을 나누는 방법 및 변환 계수들의 위치에 따른 가중치를 설정하기 위해서 DC 계수들을 별도로 처리하는 구조에 대해서 자세히 설명하기로 한다. 그 다음 3장에서는 변환 및 양자화에 관해서 자세히 설명하기로 한다. 3장에서 설명될 내용은 4x4 정수 변환 및 Intra 16x16 모드일 경우에 사용되는 DC 변환 그리고 색상 신호의 DC 계수에 쓰이는 Chroma DC 변환을 포함하고 있다. 4장에서는 전체적으로 본 논문에서 서술된 내용을 정리하고, 앞으로 연구 가능성이 높은 사항들에 대해서 간략하게 제의함으로써 끝맺고자 한다.

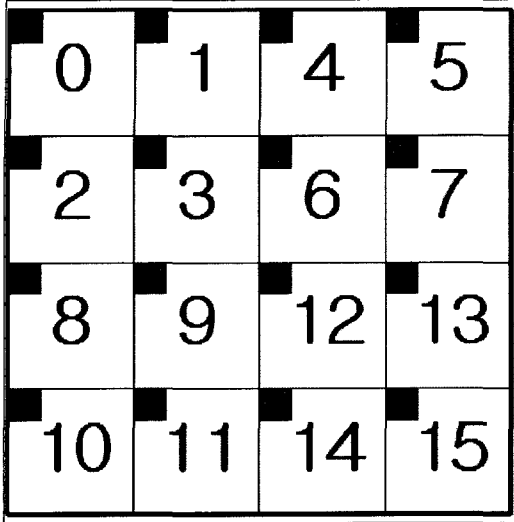


〈그림 1〉 4개의 8x8 구획으로 구성된 매크로블록 블록의 구조

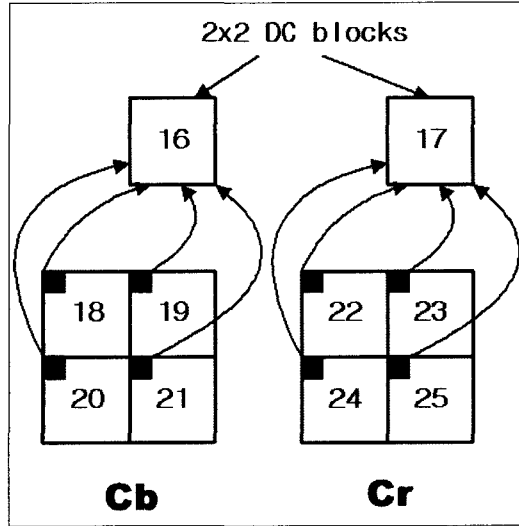
## 2. 매크로 블록의 구조

H.264 | MPEG-4 Part 10 표준에서 정의하고 있는 매크로블록(macroblock)의 크기는 그림 1에서 보여지는 바와 같이 16x16의 크기를 가지는 화소(pixel)들의 집합으로 정의된다. 그림 1에서 보여지는 매크로블록은 0~3 까지의 인덱스(index)를 가지는 4개의 8x8의 구획으로 구성되어 있는 상태를 나타낸 것이다. 0~3까지의 4개의 8x8 구획들은 변환 계수들을 부호화 할 때, 그 순서대로 부호화되는 것을 나타낸 것이다. 또한 CBFY(Coded Block Pattern for Y)를 결정하는데 있어서 각각의 8x8 구획내에 0이 아닌 변환계수들의 존재 여부에 따라서 하도록 표준안에 정의되어 있다.

〈그림 2〉는 하나의 매크로블록이 16개의 4x4 구획으로 구성되어 있음을 보여주고 있다. 〈그림 1〉에서 이미 설명한 바와 같이 하나의 매크로블록은 4개의 8x8 구획으로 나뉘어져 지정된 순서를 통해서 처리하도록 정의되어 있다. 또한 이와 마찬가지로



〈그림 2〉 16개의 4x4 구획으로 구성된 매크로블록 블록의 구조



〈그림 3〉 매크로블록의 색상 신호의 구성

하나의 8x8 구획은 4개의 4x4 구획으로 나뉘어져서 지정된 순서대로 처리되도록 정의된다. 이러한 일련의 구성은 〈그림 2〉에서 보여지고 있다. 〈그림 2〉에서는 16개의 4x4 구획의 각 DC 성분을 모아 Y라고 표시되는 4x4 구획을 다시 구성하는 것에 대해서도 나타내고 있다. 각 4x4 구획의 왼쪽 윗부분에 진하게 표시된 부분은 개념적으로 각 4x4 구획의 변환 계수들중 DC를 표시하는 부분이며, 이들 DC 계수들을 모아서 별도의 4x4 구획을 형성하는 것이 가능하게 된다. 이는 Intra 16x16 모드로 부호화 할 경우에는 DC 변환을 시행하는 것을 설명하기 위한 것으로서 Intra 16x16 DC 변환에 대해서는 2.2절에서 자세히 설명하기로 한다.

〈그림 3〉에서는 색상 신호 (chrominance signal)의 구조를 보이고 있다. 현재 H.264 | MPEG-4 Part 10 표준에서는 4:2:0 형식의 입력만을 지원하도록 정의하고 있기 때문에 하나의 매크로블록에는 두 가지의 색상 신호 Cb와 Cr의 경우에 대해서 각각 8x8의 구획만이 존재하게 된다. 이러한 구조

를 나타낸 것이 〈그림 3〉이다. 또한 〈그림 2〉의 경우에서의 유사하게 색상신호의 각 4x4 구획에서 DC 성분들을 별도로 모아서 다른 두개의 구획을 형성하게 된다. 이는 색상 신호의 DC 변환을 이용하여 압축율을 높이고자 하는 목적으로 행해지는 과정이다. 따라서 색상 신호의 경우에는 Cb와 Cr 각각의 경우에, 1개의 2x2의 크기로 구성되는 DC 구획과 4개의 4x4 크기로 이루어지는 변환 계수들의 구획들이 존재하게 된다. 따라서 하나의 매크로블록 내에 존재하는 색상 신호는 총 2개의 2x2 DC 구획과 8개의 4x4 변환 계수 구획들이 되므로 총 10개의 처리 구획들이 존재하게 된다. 이러한 일련의 구조가 〈그림 3〉에 자세하게 설명되어 있다.

이후 3장에서는 지금까지 설명한 잔류신호를 위한 4x4 정수 변환과 Intra 16x16 모드일 경우에 사용되는 DC 변환 그리고 색상 신호의 DC 변환에 사용되는 2x2 DC 변환에 대해서 자세히 설명하기로 한다.

### 3. 변환 및 양자화 (Transform & Quantization)

현재 H.264 | MPEG-4 Part 10 표준에서 정의한 변환은 기본적으로 곱셈 연산이 없이 16비트 연산만으로 수행이 가능하도록 정의되어 있다. 즉, 모든 연산을 위한 기저(kernel)는 1 및 2의 승수(power of 2) 만으로 정의가 되어있기 때문에 곱셈 연산을 하지 않고도 구현할 수 있는 구조로 설계되어 있다.

H.264 | MPEG-4 Part 10 표준의 baseline에 정의되어 있는 변환 연산은 세 가지로 나누어 질 수 있다. 첫번째가 4x4 잔류 신호 변환을 위한 정수 변환이고, 두번째는 Intra 16x16 모드로 부호화 할 경우에 4x4의 DC 신호의 변환에 사용되는 DC 변환 그리고 색상 신호(Chrominance signal)의 DC 성분의 변환에 사용되는 Chroma DC 변환이다.

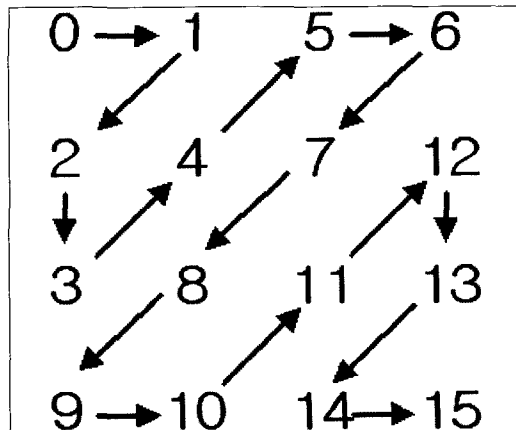
본 장에서는 각각의 변환 과정 및 정의에 대해서 자세히 설명하기로 한다.

또한 H.264 | MPEG-4 Part 10 표준에서의 양자화는 기존의 H.263이나 MPEG-4 Part 2의 표준들에서 정의된 양자화 기법과는 근본적인 구조를 달리하고 있다. H.264 | MPEG-4 Part 10 표준에서 정의된 양자화 파라미터들은 변환계수들에 직접적인 연산을 행하는 계수의 개념을 가지는 것이 아니라, 양자화를 위한 데이터의 일부분을 지칭하는 인덱스(index)의 개념을 가지는 것이 특징이라 할 수 있다. 또한 변환과 양자화가 동시에 이루어지는 특이한 구조를 내포하고 있으며, 이는 완전 복원(perfect reconstruction)이 되지 않는 H.264 | MPEG-4 Part 10 표준의 변환의 특성에 기인한 것이기도 하다. 이러한 특성에 기인하여 변환 계수들은 각 구획 내에서의 위치에 따라서 양자화되는

스케일(scale)이 달라지는 특징적인 형태를 보이게 되기도 한다. 본 장에서는 이와 같이 변환과 매우 밀접하고 연관을 가지는 양자화 과정에 대해서도 본 장에서 변환 과정의 서술과 함께 자세하게 서술하기로 한다. 3.4절에서는 휘도 신호와 색상 신호간의 양자화 파라미터의 관계에 대해서 알아보기로 한다.

#### 3.1 4x4 정수 변환

4x4 정수 변환은 Intra 및 Inter 모드에서 4x4 구획의 잔류신호(residual signal)의 압축을 위해서 사용하는 변환이다. 상기한 바와 같이 H.264 | MPEG-4 Part 10 의 모든 변환은 덧셈 및 비트 시프트(bit shift) 연산만으로 구현이 가능하기 때문에, 모든 기저는 1과 2 또는 2의 승수(power of 2) 만으로 정의가 되어 있게 된다. 기본적인 4x4 정수 변환은 <그림 4>에서 보이는 바와 같이 4x4 구획에 대해서 zig-zag 스캔을 실행하기 위한 변환계수(transform coefficients)를 만들어 내기 위해서 사용된다.



< 그림 4 > 4x4 구획의 변환 계수들에 대한 zig-zag 스캔

H.264 | MPEG-4 Part 10 표준은 여타 표준과 마찬가지로 복호기의 동작을 정의함으로써 부호기에서 성능을 높일 수 있는 여지를 남겨두고 있다. 따라서 표준안에서도 복호기에서의 역변환 연산을 정의함으로써 변환 과정을 표준화하고 있다. 이에 따른 표준의 4x4의 역변환은 아래에서 서술하는 바와 같이 정의되어 있다[1].

역변환 과정에 앞서서 복호된 4x4 계수들은 식 (1)에 의해서 스케일링 과정을 거친다. 식 (1)에서  $c_{ij}(i, j=0 \dots 3)$ 은 복호된 4x4 구획의 계수들이며  $w_{ij}(0 \dots 3)$ 는 스케일링 된 값들을 나타낸다. 또한 QP는 양자화 파라미터를 나타내며 양자화 파라미터에 대해서는 3.4절에서 보다 자세한 설명을 거치기로 한다.

$$w_{ij} = [c_{ij} \cdot R_{ij}^{(QP \% 6)}] \ll (QP / 6), \quad i, j = 0, \dots, 3 \quad (1)$$

이 때  $R_{ij}^{(m)}$ 의 값에 대해서 자세히 알아보기로 하자.  $R_{ij}^{(m)}$ 는 식 (2)에서 설명되는 것과 같은 값을 가진다.  $R_{ij}^{(m)}$ 가 가지는 값의 의미와 과정에 대해서는 3.2 절에서 자세히 다루기로 한다.

$$R_{ij}^{(m)} = \begin{cases} V_{m0} & \text{for } (i, j) \in \{(0,0), (0,2), (2,0), (2,2)\}, \\ V_{m1} & \text{for } (i, j) \in \{(1,1), (1,3), (3,1), (3,3)\}, \\ V_{m2} & \text{otherwise;} \end{cases} \quad (2a)$$

$$V = \begin{bmatrix} 10 & 16 & 13 \\ 11 & 18 & 14 \\ 13 & 20 & 16 \\ 14 & 23 & 18 \\ 16 & 25 & 20 \\ 18 & 29 & 23 \end{bmatrix} \quad (2b)$$

식 (1)의 과정을 통해서 역변환을 위한 스케일

링(scaling)된 4x4 계수들을 식 (3)과 같이 구할 수 있다. 식 (3)에서의 같이 역변환해야 할 변환 계수들이 정의되어 있을 경우에, 역변환 과정은 식 (4)~식 (5) 에서 보여지는 과정과 같이 정의되어 있다.

$$W = \begin{bmatrix} w_{00} & w_{01} & w_{02} & w_{03} \\ w_{10} & w_{11} & w_{12} & w_{13} \\ w_{20} & w_{21} & w_{22} & w_{23} \\ w_{30} & w_{31} & w_{32} & w_{33} \end{bmatrix} \quad (3)$$

a) 우선 식 (4)에서 보여지는 바와 같은 중간 변환을 시행한다.

$$z_0 = w_0 + w_2 \quad (4a)$$

$$z_1 = w_0 - w_2 \quad (4b)$$

$$z_2 = (w_1 \gg 1) - w_3 \quad (4c)$$

$$z_3 = w_1 + (w_3 \gg 1) \quad (4d)$$

b) 이후에 식 (5)에서 보여지는 것과 같은 연산을 통해서 최종적인 계수들을 얻게된다.

$$x_0 = z_0 + z_3 \quad (5a)$$

$$x_1 = z_1 + z_2 \quad (5b)$$

$$x_2 = z_1 - z_2 \quad (5c)$$

$$x_3 = z_0 - z_3 \quad (5d)$$

이러한 역변환 과정은 각 행에 대해서 먼저 실행되고 난 후, 각 열에 대해서 실행하여 최종적으로 2차원 역변환 계수들을 얻을 것을 정의하고 있다. 또한 표준안에서 각 행을 변환하는 과정에서 나 열을 변환하는 과정에서  $z_0, z_1, z_2, z_3, x_0, x_1, x_2$ , 그리고  $x_3$ 의 어떠한 값들도  $-2^{15}$ 에서  $2^{15}-1$ 의 범위를 벗어나게 하지 않을 것을 명확히 하고 있다. 이와 같이 정의함으로써 어떠한 양자화 파라미터

(QP)를 사용할 경우에도 16비트의 단위를 넘어가는 범위의 숫자를 저장해야 하는 경우를 방지하고 있다.

또한 역변환 과정에서 두 번째 단의 변환의 결과로 얻어진  $x_0, x_1, x_2, x_3$ 의 값들은 어떠한 경우에도  $-2^{15}$ 에서  $2^{15}-33$ 의 범위를 벗어나지 않도록 정의하고 있다. 이러한 제약을 두는 이유는 역변환후 최종적으로 식 (6)에서 보여지는 바와 같은 비트 이동 (bit shift)을 통하여 화소값들을 구해야 하기 때문이다.

$$X''_{ij} = [x'_{ij} + 2^5] \gg 6 \quad (6)$$

### 3.2 4x4 정수 변환의 고찰

이와 같은 4x4 정수 변환 과정을 순방향 변환과 역방향 변환의 개념으로 정립해 보면 각각 식 (7)에서 보여지는 과정과 같이 나타나게 된다.

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_{00} & x_{01} & x_{02} & x_{03} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ x_{20} & x_{21} & x_{22} & x_{23} \\ x_{30} & x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \quad (7a)$$

$$X' = \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} y'_{00} & y'_{01} & y'_{02} & y'_{03} \\ y'_{10} & y'_{11} & y'_{12} & y'_{13} \\ y'_{20} & y'_{21} & y'_{22} & y'_{23} \\ y'_{30} & y'_{31} & y'_{32} & y'_{33} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \frac{1}{2} & -\frac{1}{2} & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} \end{bmatrix} \quad (7b)$$

이제 H.264 | MPEG-4 Part 10 표준에서 정의하고 있는 변환의 특성에 대해서 살펴보기로 한다. 1차원 정수 변환 기저를 H라고 하면 H는 식 (8a)에서 보여지는 바와 같이 표현할 수 있으며, 마찬가지로 역변환 기저를 G라고 하면 G는 식 (8b)에서 보여지는 바와 같이 표현된다.

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad (8a)$$

$$G = \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} \quad (8b)$$

식 (7)에서 보여지는 H와 G를 연산하였을 때 식 (9)와 같은 결과를 얻을 수 있다. 이 때 H와 G를 기저로 하는 변환은 완전 복원(perfect reconstruction) 조건을 만족하지 않는 것을 알 수 있다[2].

$$H \cdot G = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix} \neq I \quad (9)$$

즉, 순방향 변환과 역방향 변환 기저들이 직교(orthogonal)하기는 하지만, 정규화가 되어 있지 않음으로 인해서 완전 복원을 보장하는 변환이 될 수는 없게 되는 것이다. 이러한 특성으로 인해서 양자화 과정에서 4x4 구획의 변환 계수들을 양자화 할 때, 각 위치에 따른 가중치를 달리 두어서 양자화를 실행하게 된다. 이러한 이유로 식 (2)에서와 같이  $R_j^{(m)}$ 가 결정되는 이유이다.

식 (2)에서  $V_{ml}(m=0 \dots 5, l=0 \dots 2)$ 는 3개의 열과 6개의 행을 가지는 행렬로 구성되어 있다. 이는 식 (8)에서 보듯이 H와 G를 연산하였을 때, 정규 행렬(identity matrix)의 상수배 행렬을 형성하지 않기 때문에, 이를 생성하기 위해서는 식 (10)에서 보이는 바와 같이 Q라는 가중치를 가지는 행렬을 중간에 삽입해 주어야 한다.

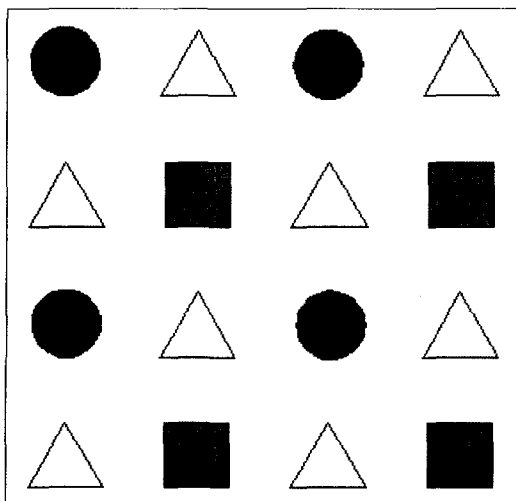
$$H \cdot Q \cdot G = 20 \cdot I \quad (10a)$$

$$Q = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \quad (10b)$$

또한 4x4 구획을 변환하기 위해서는 수직 방향과 수평 방향을 각각 한 번씩 변환해 주는 과정에서  $Q \cdot Q^T$ 의 연산을 행하게 된다. 이때 4x4 계수들에 는 각각 그림 5에서 보이는 것과 같이 위치에 따라서 서로 다른 가중치가 적용되게 된다.

<그림 5>에서 서로 다르게 표시된 위치에 있는 세 가지의 종류의 계수들은 각각 식 (2)에서 설명된 바와 같이  $V_{m\ell} (m=0 \cdot \cdot 5, \ell=0 \cdot \cdot 2)$ 의 서로 다른 세 개의 열에 해당하는 가중치로 스케일링 되게 된다. 이러한 이유로 인해서  $R_{ij}^{(m)}$ 가 변환 계수의 위치에 따라서 서로 다른 세 종류의 특성을 가지는 가중치의 값을 가지게 되는 것이다.

또한 H.264 | MPEG-4 Part 10 표준에서는 각



<그림 5> 4x4 구획 내에서의 위치에 따른 변환 계수들의 구분

양자화 단계(quantization step)간에 가중치가 12% 씩 증가하도록 정의하고 있다. 따라서 양자화 파라미터가 6의 차이를 보일때마다 2배씩 증가한 결과를 얻도록 정의되어 있음으로 인해서  $V_{m\ell}$ 에 6개의 값만이 정의되어 있으며, QP의 값의 증가에 따라서 6으로 나눈 값과 6으로 모듈로(modulo) 연산한 값만을 가지고 반복적으로 양자화를 실행하게 되는 것이다.

### 3.3 Intra 16x16 DC 변환

Intra 16x16 DC 변환은 상기한 바와 같이 하다마드 변환(Hadamard transform)을 사용한다. Intra 16x16 모드로 부호화된 매크로블록(macroblock)은 4x4 구획으로 나뉘어져서 변환되게 된다. 이 후 다시 16개의 DC 계수들을 모아서 DC들로만 이루어진 4x4 구획을 구성하여 이를 다시 하다마드 변환하여 부호화하는 것을 정의하고 있다. 따라서 Intra 16x16 DC 역변환은 4x4 변환 이전에 행해져야 할 과정이 된다.

<그림 2>에서 어떠한 Intra 16x16 모드로 결정되었을 경우에는, 16개의 4x4 구획들을 정수 변환하여 얻은 변환 계수(transform coefficients)들 중에서 각 구획의 DC 성분들만을 수집하여 별도의 4x4 구획을 형성하는 과정을 이미 설명한 바 있다. 이는 본 절에서 설명하고자 하는 Intra DC 변환을 행하기 위한 과정의 일환이다. 이와 같이 어떠한 매크로블록이 Intra 16x16 모드일 경우에는 일반적으로 입력 영상의 평탄한 부위에 해당하는 확률이 매우 높은 데서 기인한다. 이러한 경우 일반적으로 DC 성분들 사이에도 서로 상당한 양의 상관관계가 존재할 확률이 매우 높으며, 이는 바로 DC 성분들을 모아서 다시 변환을 행하였을 경우에 에너지 집중(energy compaction)의 효율 또한 기대할 수 있다

는데서 제안된 것이다.

또한 Intra 16x16 DC 변환은 하다마드 변환을 통해서 이루어진다는 것을 이미 밝힌 바 있으며, 이는 식 (11)에서 보여지는 바와 같이 나타나게 된다.

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad (11)$$

역변환이 행해진 후에는 스케일링 (scaling)을 행하여 결과로 얻어진 계수들의 범위를 유효하게 해주어야 한다. 이러한 스케일링은 식 (12)에서 보여지는 바와 같이 행해지게 된다.

- a) 양자화 파라미터의 크기가 12와 같거나 12보다 클 경우

$$DC_{ij} = [F_{ij} \cdot R_{00}^{(QP * 6)}] \ll (QP/6 - 2), \quad i, j = 0, \dots, 3 \quad (12a)$$

- b) 양자화 파라미터의 크기가 12보다 작을 경우

$$DC_{ij} = [F_{ij} \cdot R_{00}^{(QP * 6)} + 2^{1-QP/6}] \gg (2 - QP/6), \quad i, j = 0, \dots, 3 \quad (12b)$$

H.264 | MPEG-4 Part 10 표준에서는 결과로 얻어진 어떠한  $DC_{ij}$  값도  $-2^{15}$ 에서  $2^{15}-1$ 의 범위를 벗어나지 못하도록 제약하고 있다.  $R_{ij}^{(m)}$ 의 값이 6을 반복적으로 나타내는 이유는 이미 3.2절에서 설명한 바와 같다.

### 3.4 Chroma DC 변환

Chroma DC 변환 또한 Intra 16x16 DC 변환과 유사한 개념을 가진다. Chroma DC 변환도 2x2의

하다마드 변환을 사용한다. 이러한 이유는 현재 H.264 | MPEG-4 Part 10 표준에서는 4:2:0 형식의 입력 영상만을 지원하도록 정의하고 있기 때문에 휘도 성분의 신호 (luminance signal)에서는 하나의 매크로블록 내에 16개의 4x4 구획들이 존재하게 되지만, 색상 신호 (chrominance signal)의 경우에는 하나의 매크로블록 내에 각각 4개씩의 4x4 구획들만이 존재하기 때문이다. 따라서 DC 변환은 이들 4개의 구획에서 모여진 DC 성분을 추출하여 하나의 구획을 형성할 경우에 2x2 구획을 형성하게 된다. 따라서 2x2 하다마드 변환이 적용되게 된다. 이와 같이 매크로블록의 색상 신호가 2개의 2x2 크기의 DC 성분을 가지는 구획과 8개의 4x4 크기의 AC 성분을 가지는 구획으로 구성되는 일련의 과정은 <그림 3>에서 자세히 설명한 바 있다.

Chroma DC 변환 과정 또한 Intra DC 변환과 마찬가지로 하다마드 변환을 사용한다. 다만 Chroma DC 변환은 대상 구획의 크기가 4x4의 크기가 아니라 2x2의 크기를 가지기 때문에 Intra DC 변환의 경우와는 조금 다르게 된다. 색상 신호의 4개의 4x4 변환 구획에서 수집된 4개의 DC 변환계수를 각각  $C_{00}$ ,  $C_{01}$ ,  $C_{10}$ ,  $C_{11}$ 라고 할 때, Chroma DC 변환  $F$ 의 과정은 식 (13)에서와 같이 정의된다.

$$F = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (13)$$

또한 식 (12)에서 이미 설명된 바와 같이 역변환의 과정을 거친 후에는 식 (14)에서 설명하는 바와 같은 스케일링 (scaling) 과정을 거치는 것 또한 Intra DC 변환의 과정과 매우 유사하다.



a) 양자화 파라미터의 크기가 6과 같거나 6보다 클 경우

$$DC_{ij} = [F_{ij} \cdot R_{00}^{(QP \% 6)}] \ll (QP / 6 - 1), i, j = 0, \dots, 3 \quad (14a)$$

b) 양자화 파라미터의 크기가 6보다 작을 경우

$$DC_{ij} = [F_{ij} \cdot R_{00}^{(QP \% 6)}] \gg 1, i, j = 0, \dots, 3 \quad (14b)$$

H.264 | MPEG-4 Part 10 표준에서는 결과로 얻어진 어떠한  $DC_{ij}$  값도  $-2^{15}$ 에서  $2^{15}-1$ 의 범위를 벗어나지 못하도록 제약하고 있다.  $R_{ij}^{(m)}$ 의 값의 설명에 대해서는 이미 3.2절에서 설명한 바 있으므로 생략하기로 한다.

### 3.5 휘도 신호와 색상 신호의 양자화 파라미터 관계

본 절에서는 변환 기법과 매우 밀접한 관계를 가지고 있는 양자화 파라미터에 대해서 자세히 설명하는 것을 그 목적으로 한다. H.264 | MPEG-4 Part 10 표준에서는 0~51까지의 52단계의 양자화 단계(quantization step)을 정의하고 있다. 색상 신호는 그 특성상 휘도 신호에 비해서 그 복잡도가 덜하므로, 휘도 신호에 지정된 양자화 파라미터(quantization parameter)를 똑같이 사용하지 않는다. 이러한 기법은 H.263의 경우에서도 찾아볼 수 있는 기법이다. 이러한 이유로 인해서 표준에서 30~51까지의 범위내에서 정의된 양자화 단계는 휘도 신호와 색상 신호가 각각 다르게 지정되어 있다. <표 1>에서는 30~51의 범위에서 지정된 휘도 신호의 양자화 파라미터에 해당하는 색상 신호의 양자화 파라미터의 관계를 나타내고 있다. 양자화 파라미터가 0~29까지의 값을 가지는 경우에는 휘도 신호와 색상 신호가 같은 양자화 파라미터 값을

< 표 1 > 휘도 신호와 색상신호의 양자화 파라미터의 관계

30	31	32	33	34	35	36	37	38	39	40
29	30	31	32	32	33	34	34	35	35	36
41	42	43	44	45	46	47	48	49	50	51
36	37	37	37	38	38	38	39	39	39	39

사용하도록 정의되어 있다.

## 4. 결론

본 논문을 통해서 상세히 서술한 바와 같이 H.263, MPEG-4 등과는 달리 H.264 | MPEG-4 Part 10 표준은 정수 변환(integer transform)을 채택함으로써 기존의 표준에 의거할 경우에 이산 여산 변환(Discrete Cosine Transform)에 의해서 발생할 여지가 있었던 오차(mismatch)를 근본적으로 제거한 것이 가장 큰 특징이라 할 수 있겠다. H.264 | MPEG-4 Part 10에서는 정수 연산 및 비트 이동(bit shift) 연산만으로 변환을 정의함으로써 인해서 디지털 시스템에서 구현 시에 연산으로 인한 오차가 날 수 있는 여지를 없앴다.

또한 4x4 정수 변환 연산은 8비트 입력을 가정하였을 때 16비트 기반의 시스템에서 직접적인 연산만으로 구현이 가능하도록 정의되어 있다는 점 또한 기존의 표준들이 가지지 못했던 장점으로 꼽을 수 있다.

H.264 | MPEG-4 Part 10 표준에서 정의된 변환의 또 다른 특징은 양자화 과정과 매우 밀접한 연관을 가지고 실행되고 있다는 점이다. 변환과 이에 따른 비트 이동 연산시에 정수 변환이 가지는 특성으로 말미암아, 이산 여산 변환(Discrete Cosine Transform)의 경우에는 볼 수 없었던 변환 계수의 위치에 따른 가중치를 설정하는 특성 또한 본 논문

의 상당 부분에 걸쳐서 자세히 설명된 바 있다.

현재 H.264 | MPEG-4 Part 10 표준은 위원회 최종안(FCD : Final Committee Draft)의 과정에 있다. 현 상황에서 변환 및 양자화에 관한 근원적인

연구는 많이 진행된 상태이나, 양자화 단계에 따른 응용에서의 성능 및 구현에 관한 여러 가지 분야의 연구는 학술적, 산업적으로 많은 진전을 요구하고 있다고 여겨진다.

● 참고 문헌 ●

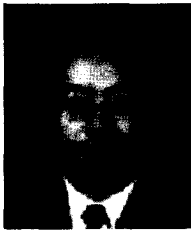
(1) Thomas Wiegand & Gary Sullivan, "Joint Final Committee Draft (JFCD) of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)," July, 2002.  
 (2) Henrique Malvar, "Low-Complexity Length-4 Transform and Quantization with 16-bit Arithmetic," VCEG-N44 Contribution, Sept. 2001.  
 (3) Louis Kerofsky, "H.26L Transform/Quantization Complexity Reduction Ad Hoc Report," VCEG-009 Contribution, Nov. 2001.

필자 소개



김철우

-1987년 3월~1991년 2월 : 서울대학교 전자공학과 학사  
 -1991년 3월~1993년 2월 : 서울대학교 전자공학과 석사  
 -1993년 3월~1998년 2월 : 서울대학교 전자공학과 박사  
 -1998년 3월~2000년 3월 : LG 전자 DM연구소 선임연구원  
 -2000년 3월~2000년 12월 : ACS Technology 책임연구원  
 -2001년 1월~현재 : (주)엠큐브웍스 기술이사  
 -주관심분야 : H.263 및 H.26L, MPEG-4 비디오 표준화 및 멀티미디어 통신 시스템



설원희

-1984년 2월 : 서울대학교 제어계측공학과 학사  
 -1986년 5월 : Michigan 대학교 컴퓨터 공학 석사  
 -1993년 5월 : Purdue 대학교 컴퓨터 공학 박사  
 -1993년 4월 : Miami 대학교 전기 컴퓨터공학과 연구 조교수  
 -1995년 3월 : GE Medical Systems 선임연구원  
 -1997년 9월 : MCC(Microelectronics and Computer Technology Corporation) 수석연구원  
 -2000년 1월 : SK Telecom 정보기술원 IT 인프라개발그룹장  
 -2000년 12월 : SK Telecom Platform 연구원 원장  
 -주관심분야 : 이동통신 시스템, H.263, H.26L, MPEG-4 비디오 표준화 및 3GPP 멀티미디어 표준