

# JPEG 파일 크기를 제어하기 위한 DPCM 기반의 영상 사전 분석기와 양자화 방법

신선영, \*고혁진, \*\*박현상, 전병우  
성균관대학교 전자공학과

\*삼성전자 SYS.LSI 사업부, Mobile 개발실, Image 팀  
\*\*공주대학교 전기전자공학부

## DPCM-Based Image Pre-Analyzer and Quantization Method for Controlling the JPEG File Size

Sun-Young Shin, \*Hyuk-Jin Go, \*\*Hyun-Sang Park, Byeung-Woo Jeon  
Department of Electronic Engineering  
Sungkyunkwan University

E-mail : [cathyshin@samsung.com](mailto:cathyshin@samsung.com), [hyukjin.ko@samsung.com](mailto:hyukjin.ko@samsung.com),  
[vandamm@kongju.ac.kr](mailto:vandamm@kongju.ac.kr), [bjeon@yurim.skku.ac.kr](mailto:bjeon@yurim.skku.ac.kr)

### Abstract

In this paper, we present a new JPEG (Joint Photograph Experts Group) compression architecture which compresses still image into fixed size of bitstream to use restricted system memory efficiently. The size of bitstream is determined by the complexity of image and the quantization table. But the quantization table is set in advance the complexity of image is the essential factor. Therefore the size of bitstream for high complexity image is large and the size for low complexity image is small. This means that the management of restricted system memory is difficult.

The proposed JPEG encoder estimates the size of bitstream using the correlation between consecutive frames and selects the quantization table suited to the complexity of image. This makes efficient use of system memory.

### I. 서론

영상을 촬상하는 경우, 사용자는 액정 화면과 같은 디스플레이 장치에 디스플레이되는 동영상을 관찰하다가 촬상하고자 하는 영상이 나타날 때 촬상하며, 이 때, 촬상된 영상은 소정의 영상 압축

방법에 의해 압축되어 파일 형태로 저장된다. 정지영상 압축 형식은 GIF(Graphics Interchange Format) 형식, JPEG 형식[1] 등이 대표적이며, 이 중 동영상 표준으로 널리 사용되는 MPEG(Moving Picture Experts Group)[2]과 연산 구조가 동등한 JPEG 형식이 널리 이용된다.

2001년 국내에 소개되기 시작한 카메라폰은 Mega 화소급 고급 이미지 센서가 출시되면서[4], 정지 영상 압축 위주의 단순 기능으로부터 동영상 압축을 포괄하는 고기능, 다기능 제품으로 진화를 거듭하고 있다. 이와 더불어 고급 기능을 필요로 하지 않는 시장의 요구에 대응하기 위해서 단순 기능의 카메라폰의 수요도 동반 상승하고 있으며, 전세계적인 시장의 규모로 볼 때는 오히려 단순 기능의 카메라폰이 오히려 높은 시장 점유율을 차지하고 있다[3].

단순 기능의 카메라폰의 경우 개발 단가를 낮추기 위해서 가급적 적은 용량의 메모리를 사용해야 한다. 따라서 정지 영상을 압축한 후, 이를 저장하기 위한 메모리 공간이 제한적이기 때문에 메모리를 효과적으로 사용할 수 있는 정지 영상 구조가 필요하게 된다. 종래의 영상압축 구조에서는 촬상되는 영상의 복잡도에

따라 압축된 영상의 크기가 달라지기 때문에, 제한된 용량의 메모리에 오버플로(overflow)가 발생하지 않도록 다음과 같은 방법을 사용해왔다. 첫째, 영상의 복잡도에 상관없이 충분히 높은 압축률로 압축하는 것이다. 이는 구현이 용이하나 화질의 열화가 심하다는 단점이 있다. 또 다른 방법은 오버플로가 발생할 경우에는, 오버플로가 발생하지 않을 때까지 활상을 반복하는 것이다. 이 경우 고화질을 유지할 수 있으나, 프로세서의 JPEG 엔진에의 개입이 필요하고, 활상에 요구되는 실질적인 시간이 지나치게 길어지는 문제가 야기되어 큰 전력소모나 연속 촬영의 구조적 어려움 등의 부작용을 초래한다[7].

본 논문에서는 활상 영상을 이용하여 예측을 하는 경우[5][6] 시간적 지연을 막기 위해 활상 이전에 미리 예측하여 적절한 양자화 테이블로 압축함으로써 일정한 크기로 압축된 영상의 크기를 제어할 수 있는 JPEG 압축 구조를 제안한다. 제안한 구조에서는 활상하고자 하는 영상 프레임과 직전 영상 프레임과의 높은 상관관계를 이용하여, 직전 프레임에 대한 압축 영상의 크기를 예측하고, 이를 기반으로 최적의 양자화 테이블을 선택하여 현재 영상을 부호화하도록 한다. 예측기법으로는 DPCM 을 사용하기 때문에 기존의 예측기법에 비해[5][6][9] 계산량이 적어서 하드웨어 설계가 용이하다.

본론에서는 기존 JPEG 압축 구조 및 문제점을 분석하고, 일정한 크기의 JPEG 압축 파일을 얻기 위한 예측방법과 그 구조에 대해서 기술한다.

## II. 사전 영상 분석 및 양자화 방법

### 2.1. 기존 정지 영상 압축기 구조

기존의 이산 여현 변환 (DCT:Discrete Cosine Transform) 기반의 JPEG 압축기는 그림 1 과 같은 순서로 데이터를 처리한다.

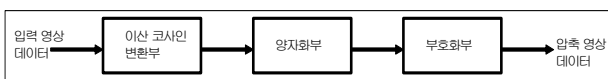


그림 1. 정지 영상 압축 구조

이와 같은 구조에서 압축은 MCU (Minimum Coded Unit) 단위로 순차적으로 수행된다. 이산 여현 변환부는 사용자에게 의해 활상된 영상의 영상 데이터를 이산 여현

변환하여 출력한다. 양자화(Quantization)부는 사전에 결정된 양자화 테이블에 따라 이산 여현 변환된 결과를 양자화한다. 부호화부(Variable Length Code)는 양자화된 결과를 Huffman 부호화하여 압축영상 데이터를 출력한다.

일반적으로 JPEG 에서 사용되는 기본 양자화 테이블 (QTable0)은 다음과 같다[1].

|    |    |    |    |     |     |     |     |
|----|----|----|----|-----|-----|-----|-----|
| 16 | 11 | 10 | 16 | 24  | 40  | 51  | 61  |
| 12 | 12 | 14 | 19 | 26  | 58  | 60  | 55  |
| 14 | 13 | 16 | 24 | 40  | 57  | 69  | 56  |
| 14 | 17 | 22 | 29 | 51  | 87  | 80  | 62  |
| 18 | 22 | 37 | 56 | 68  | 109 | 103 | 77  |
| 24 | 35 | 55 | 64 | 81  | 104 | 113 | 92  |
| 49 | 64 | 78 | 87 | 103 | 121 | 120 | 101 |
| 72 | 92 | 95 | 98 | 112 | 100 | 103 | 99  |

그림 2. 기본 양자화 테이블

기본 양자화 테이블의 양자화 계수들을 8 로 나누어 생성된 양자화 테이블 (QTable0/8)로 VGA 영상 100 장에 대해서 압축을 수행한 결과 각 프레임의 압축 영상 크기는 그림 3 과 같이 영상의 복잡도에 따라 다양한 크기의 압축 영상을 얻게 된다.

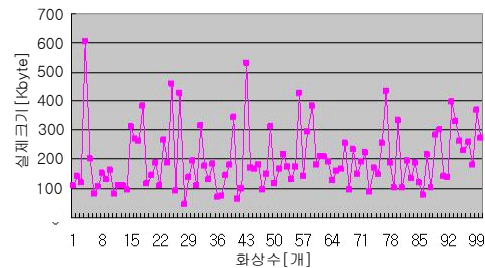


그림 3. 기본 테이블에 의한 JPEG 압축 파일 크기

### 2.2. 제안한 정지 영상 압축기 구조

상기 문제를 해결하기 위해 본 논문에서 제안하는 JPEG 압축기는 그림 4 와 같이 구성된다.

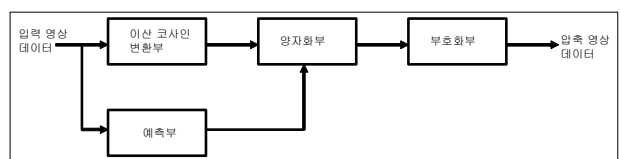


그림 4. 파일 크기 제어를 위한 정지 영상 압축 구조

예측부는 활상된 영상의 이전 프레임의 수평 방향의 고주파 성분의 양과 수직 방향의 고주파 성분의 양에 의거하여 활상된 영상에 대한 압축된 영상의 크기를 예측한다. 양자화부는 예측값에 따라 소정의 양자화 테이블을 선택하고, 선택된 양자화 테이블을 이용하여 이산 여현 변환된 영상 데이터를 양자화한다. 부호화부는 상기 양자화된 영상 데이터를 Huffman 부호화한다.

### 2.3. DPCM 기반의 영상 사전 분석기

그림 5 는 예측부의 구조 및 동작을 보여준다.

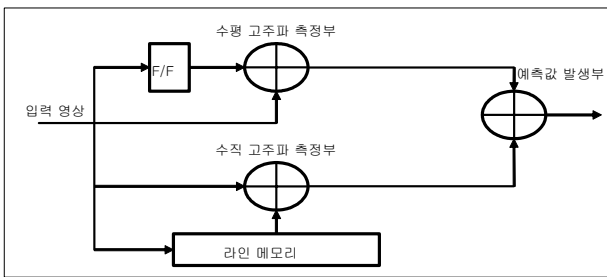


그림 5. 예측부 블록 다이어그램

예측부는 수평 고주파 측정부, 수직 고주파 측정부, 및 예측값 발생부로 구성된다. 고주파 성분은 활상된 영상의 복잡도를 나타내는 것으로서, 이웃하는 화소 간의 휘도값의 차이의 절대값을 합산함으로써 구한다.

MxN 해상도의 이전에 프레임에 대해  $I(i,j)$  ( $1 \leq i \leq M, 1 \leq j \leq N$ ) 가 수평방향으로  $i$  번째, 수직방향으로  $j$  번째에 위치한 휘도성분이라고 하자. 이때  $j$  번째 열의 고주파 성분 값( $H_j$ )는 다음과 계산된다.

$$H_j = \sum_{i=1}^{M-1} |I(i,j) - I(i+1,j)|, \quad 1 \leq j \leq N$$

동일한 방법으로 이전 프레임의 N 개의 열에 대해 고주파 성분 값을 측정하여 누산함으로써 최종 수평 방향의 고주파 성분의 양, HSUM, 을 계산한다.

$$HSUM = \sum_{j=1}^N H_j$$

수직 고주파 측정부에서 각 행의 고주파 성분의 양을 구하는 방법은 수평 고주파 측정부와 동일한 방법으로 라인 메모리를 이용하여 이웃하는 두 열 간에 동일한 수평 위치에 있는 화소에 대해 DPCM 값을 구하고, M 개의 행의 고주파 성분 값을 누산함으로써 최종 수직 방향의 고주파 성분, VSUM, 을 구한다.

$$V_i = \sum_{j=1}^{N-1} |I(i,j) - I(i,j+1)|, \quad 1 \leq i \leq M$$

$$VSUM = \sum_{i=1}^M V_i$$

예측값 발생부는 HSUM 과 VSUM 을 합산하여 다음과 같이 예측값(EV)을 발생한다.

$$EV = HSUM + VSUM$$

그림 6 은 예측값과 실제 압축된 영상의 크기와의 관계를 나타내었다. 실험을 위한 영상은 VGA 해상도의 영상 66 장을 사용하여 측정하였으며, 표의 가로축은 예측부에서 측정된 예측값을 나타내고 세로축은 실제 압축된 영상의 크기를 나타낸다.

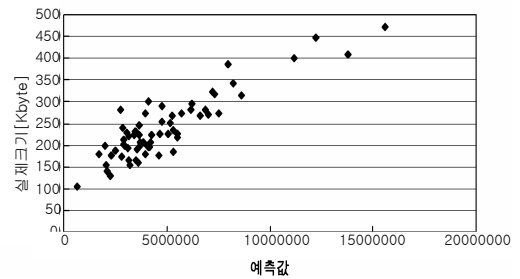


그림 6. 예측값과 실제 압축 영상 크기

그림 6 에서 예측값과 실제 압축된 영상의 크기는 큰 상관관계를 가지는 것을 알 수 있다. 즉, 예측값이 측정되면 측정된 예측값에 근거하여 압축된 영상의 크기를 예측할 수 있음을 알 수 있다. 데이터 분석 툴 Minitab 을 사용하여 상관관계를 측정한 결과 0.894 의 상관관계를 갖는 것으로 나타났다.

### 2.4. 타겟 압축 영상을 위한 양자화 테이블 선택

본 논문에서는 VGA 해상도 영상에 대해 타겟(Target)압축 영상 데이터를 128 Kbytes 로 만드는 10 개의 양자화 테이블을 그림 7 과 같이 구성하였다. 각 테이블은

- QTable1 : QTable0 / 16
- QTable2 : (3\*QTable0) / 32
- QTab:e3 : (3\*QTable0) / 16
- QTable4 : QTable0 / 4
- QTable5 : (5\*QTable0) / 32
- QTable6 : (5\*QTable0) / 16
- QTable7 : (3\*QTable0) / 8

- QTable8 : (4\*QTable0) / 8
- QTable9 : (5\*QTable0) / 8
- QTable10 : (7\*QTable0) / 8

의 계산식에 의해 생성한다. 예측값과 압축 영상 크기가 비례하므로, 4 개의 상위비트가 클수록 큰 폭으로 양자화가 수행되는 하위 양자화 테이블을 선택하여 큰 폭으로 압축이 수행되도록 하고, 4 개의 상위 비트가 작을수록 작은 폭으로 양자화가 수행되는 상위 양자화 테이블을 선택하여 작은 폭으로 압축이 수행되도록 한다. 즉, 예측값의 4 개의 최상위 비트가 0000 이면 양자화부는 Qtable1 을 선택하고, 1001 또는 그 이상이면 가장 큰 폭으로 양자화가 수행되는 Qtable10 을 선택한다.

|      |           |
|------|-----------|
| 0000 | Qtable 1  |
| 0001 | Qtable 2  |
| 0010 | Qtable 3  |
| 0011 | Qtable 4  |
| 0100 | Qtable 5  |
| 0101 | Qtable 6  |
| 0110 | Qtable 7  |
| 0111 | Qtable 8  |
| 1000 | Qtable 9  |
| 1001 | Qtable 10 |

그림 7. 양자화 테이블

### III. 결론

그림 8 는 VGA 해상도 영상 100 장에 대해 제안한 예측방법에 의해 JPEG 압축을 수행한 결과이다.

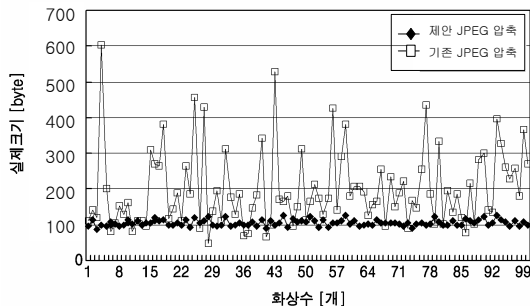


그림 8. 압축 영상 파일 크기

모든 영상이 128Kbytes 에 근접한 압축 영상을 출력함을 알 수 있다. 이상으로 제안하는 JPEG 압축기는 사람의 눈에 덜 민감한 고주파 성분을 많이 구비하는 영상에 대해서는 양자화를 더 크게 하여 압축을 수행하여 압축된 영상의 크기가 일정하도록 하므로, 효율적으로 메모리를 관리할 수 있도록 한다.

### 참고문헌

- [1] ITU CCITT Recommendation T.80 | ISO/IEC 10918-1 JPEG “Digital Compression Coding of Continuous-tone Still Image,” 1994.
- [2] Draft ISO/IEC MPEG-4 Video Verification Model Ver. 6.0, “Coding of Moving Pictures and Associated Audio Information,” Feb. 1996.
- [3] [http://www.hackersnews.org/data/2003/10/1003\\_16.html](http://www.hackersnews.org/data/2003/10/1003_16.html)
- [4] UXGA data sheet: 1/3.2” UXGA CMOS Image Sensor with an Embedded Image Signal Processor, Samsung Electronics Company, Aug 2004.
- [5] Y. Shoham and A. Gersho, “Efficient bit allocation for an arbitrary set of quantizers,” *IEEE Trans. on Acoust., Speech, Signal Processing*, Vol. 36, pp. 1445–1453, Sept. 1988.
- [6] Wook Joong Kim, Jong Won Yi and Seong Dae Kim, “A bit allocation method based on picture activity for still image coding,” *IEEE Trans. on Image Processing*, Vol.1.8, No 7, 1999.
- [7] Nakagawa et al., “DCT-based still image compression ICs with bit-rate control,” *IEEE Trans. on Image Processing*, Vol. 38, pp 711-717, Aug. 1992.
- [8] B. Arcangelo, and M. Massimo, “JPEG compression factor control: a new algorithm,” *2001.ICCE. International Conference Trans. on Consumer Electronics*, pp 206-207, June 2001.
- [9] Supangkat et al., “Quantity control for Jpeg image data compression using fuzzy logic algorithm,” *IEEE Trans. on Consumer Electronics*, Vol.1.41, Nov. 1995.