

# 인터넷 화상 회의를 위한 H.263 화상 부호화기의 속도 개선

이수익, 장주욱

서강대학교 전자 공학과

## Enhancement in Processing Speed of H.263 Encoder for Video Conferencing on Internet

Soo-Ick Lee Ju-wook Jang

Dept. of Electronic Engineering, Sogang Univ.

### 요약

실시간 인터넷 화상 회의 시스템을 위한 H.263 동화상 부호화 소프트웨어의 처리 속도를 개선하는 기법을 개발하였다. 화상 회의 영상의 특성과 H.263 프로토콜의 구조를 분석하여 데이터의 압축률이나 화질에 미치는 영향을 최소화하면서 처리 시간을 크게 단축하는 기법을 개발하였다. 처리 속도의 개선을 위해서 양자화 계수, 움직임 검출 연산, DCT 변환 연산등의 상호 연관성을 분석하여, 배경 인식이나 움직임 조기 검출, 연산 결과를 미리 예측하여 연산을 중단하는 기법등을 개발, 적용하므로써 압축률과 화질에 미치는 영향을 최소화하면서 계산량을 크게 줄이도록 하였다.

### 1. 서론

본 연구에서는 인터넷상의 개인용 컴퓨터 시스템에서 H.263이 가지는 높은 압축률을 유지하면서 자연스러운 화상 회의가 가능한 처리 속도를 가지는 부호화기 구축을 위한 기법들을 제안하였다. 이러한 기법들은 H.263 부호화 프로토콜의 구조상 가장 많은 시간을 필요로 하는 움직임 검출 과정에 대해 화상 회의 영상의 통계적 특성을 고려하여 화질과 압축률에 미치는 영향을 최소화하면서 처리 시간을 크게 단축하도록 한 것이다. 처리 속도의 개선을 위해서 양자화 계수, 움직임 검출 연산, DCT 변환 연산등의 상호 연관성을 분석하여, 배경 인식이나 움직임 조기 검출 기법등을 개발, 적용하므로써 압축률과 화질에 미치는 영향을 최소화하면서 계산량을 크게 줄이도록 하였다. 펜티엄급 PC에서 다양한 특성을 가진 벤치마크 동화상들에 대해 실험한 결과 기존의 Telenor 동화상 부호화 소프트웨어에 비해 처리 속도가 크게 향상됨을 보였다.

### 2. H.263 화상 부호화기의 구조

그림 1은 H.263 화상 부호화기의 구조를 보인 것이다. H.263은 그림 1과 같이 MC(Motion Compensation)과 DCT(Discrete Cosine Transform)을 기반으로 영상의 시간적 공간적 중복(Redundancy)을 제거 하므로써 정보 표현을 위한 데이터량을 감소시키도록 해준다. H.263의 이러한 데이터 압축은 16X16의 매크로 블록과 8X8의 블록을 기반으로 하게 되며, 16X16 매크로 블록은 시간적 중복을 고려한 움직임 벡터 검출 연산의 기본 단위가 되며, 8x8 블록은 공간적 중복을 고려한 DCT 연산의 기본 단위가 된다.

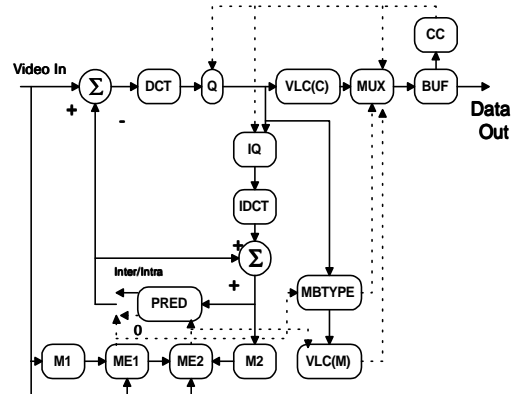
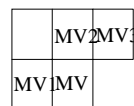
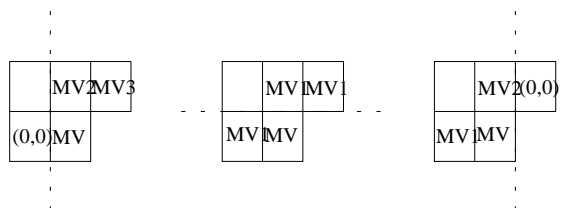


그림 1 H.263 부호화기의 구조



MV : Current motion vector  
 MV1: Previous motion vector  
 MV2: Above motion vector  
 MV3: Above right motion vector



- - - - : Picture or GOB border

그림 2 움직임 벡터 예측

움직임 벡터는 또한 그림 2와 같이 매크로 블록의 변환 자체만으로 부호화하지 않고, 주변 매크로 블록의 움직임도 고려한 예측 벡터의 차를 부호화하게 된다.

이러한 차를 부호화하는 것은 인접 매크로 블록간의 상호 관계도를 이용하여 압축율을 더 높이기 위해 사용된다.

### 3. 화상 회의 영상 데이터의 특성

화상 회의 영상 데이터의 특성으로 정성적으로 고찰되는 첫 번째 특징은 화상 회의의 특성상 화상의 상당 부분이 배경 부분으로서 연속되는 프레임상에서 거의 변화가 없다는 점을 들 수 있다. 두 번째로 움직이는 피사체의 경우는 적은 수의 물체가 한 두 가지 방향으로 움직이는 경우가 대부분이라는 것이다.

이러한 특성을 정량적으로 고찰하기 위해 몇 가지 벤치마크 동화상 샘플에 대한 실험을 실시하였다. 고려된 동화상 샘플로서는 앞서 기술한 화상 회의 영상의 특성을 매우 잘 반영한 것과 실험의 완전성에 대한 고려를 하기 위해 다소 그러한 성격이 벗어난 것들에 대한 실험을 하였다.

첫 번째로 고려한 동화상인 CLAIRE 샘플은 여자 아나운서가 일반적인 화상 회의의 양식대로 입술과 머리 부분등을 주기적으로 움직이는 것으로서 배경은 단색으로 약간의 조명 변화를 제외하고는 거의 변화가 없는 것이다. 두 번째로 고려한 CARPHONE 샘플은 남자가 움직이는 차안에서 화상 전화를 하고 있는 동화상으로 오른쪽의 창문을 통해 풍경이 차가 움직이는 반대 방향으로 상당히 빠르게 움직이는 특징을 가진 것이다. 세 번째로 고려한 FOREMAN 샘플은 남자가 상당히 동적으로 움직이는 동시에 카메라가 같이 움직여서 배경이 지속적으로 변하는 동화상으로 카메라가 90도 이상 전환되는 부분도 포함되어 있다.

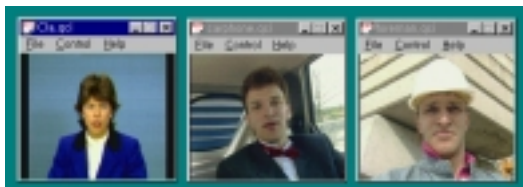


그림 3 벤치마크 동화상 샘플

실험은 각각의 샘플 동화상을 Telenor H.263 부호화기로 부호화 했을 때 나타나는 움직임 예측 벡터와 실제 움직임 벡터의 차의 분포를 조사하는 방법으로 이루어졌다. 각각의 샘플은 176X144의 해상도를 가지는 QCIF 형식의 300프레임씩으로 이루어져 있다.

그림 4에서 실험의 결과를 보였다. 실험 결과에서 가장 주목되는 특성은 움직임 예측 벡터와 실제 벡터와의 차이가 전혀 없던 매크로 블록이 상당한 부분을 차지한다는 것이다. 특히 화상회의의 전형적 특성을 가지는 CLAIRE 영상 샘플의 경우는 90% 이상이 이에 해당함을 알 수 있다. 이러한 결과는 주변 움직임 벡터를 참조하여 예측되어 지는 움직임 예측 벡터가 실제 움직임 벡터와 정확히 일치하는 경우가 상당히 많다는 것을 의미한다. 이것은 특히 화상 회의 영상 데이터처럼 배경이 상당 부분을 차지하고, 움직이는 피사체가 한 두 개의 방향으로만 움직이는 경우에 현저하게 적용됨을 알 수 있다.

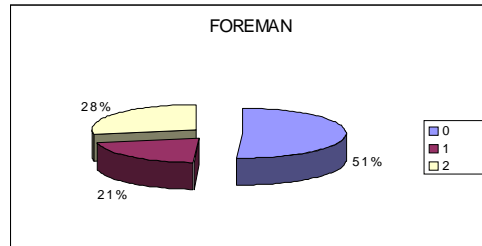
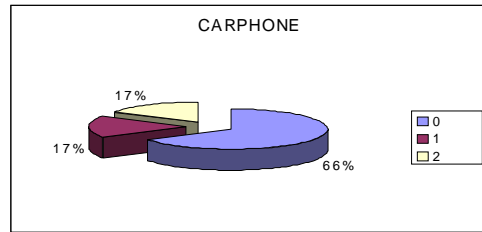
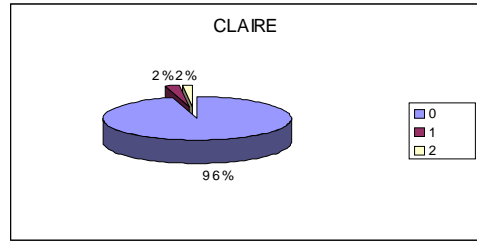


그림 4 각 동화상 샘플의 움직임 예측 오류 분포

### 4. 움직임 검색 속도의 개선

H.263 부호기의 처리 속도를 개선하기 전의 기존 기법들을 사용할 때 각각의 기능 부분에서 소요되는 시간의 분포는 그림 5와 같다.

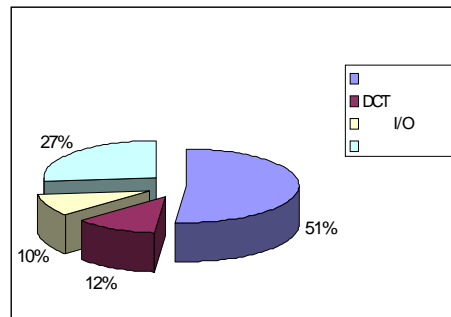


그림 5 개선전 처리 시간 분포

그림 5에서와 같이 H.263 부호화기에서 움직임 예측에 소요되는 시간이 다른 부분보다 상대적으로 상당히 큰 부분을 차지하고 있음을 알 수 있다. 이것은 기존에 구현된 H.263 부호화기가 움직임 예측의 정확도를 최대화하기 위해 전역 탐색 기법을 사용하고 있기 때문이다. 전역 탐색 방법을 사용할 경우 검색 영역 전체에 대해 SAD(Sum of Absolute Difference)들을 모두 계산해야 하므로 움직임 벡터 하나를 계산하는데 상당한 계산량을 요구하게 된다. 그러나 이러한 SAD 값의 통계적 분포를 고려하면 이러한 계산 과정을 제거할 여지가 있음을 알 수 있다. 그림 6에 일반적인 움직임 벡터 검색시의 SAD 값 분포를 보였다.

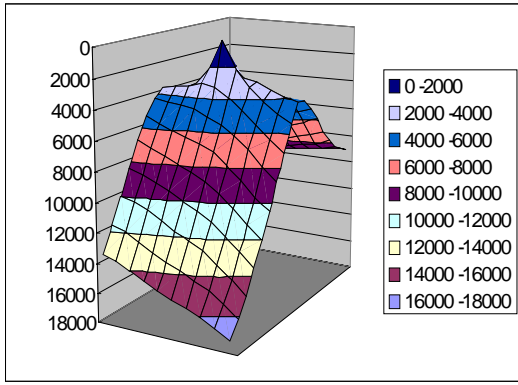


그림 6 움직임 벡터 검색시 SAD값(반전)의 일반적인 분포

SAD 값의 분포를 보면 Best-match 점을 중심으로 산 모양(SAD 값의 반전 분포)을 하고 있음을 알 수 있다. 이러한 분포는 다소 예외적인 경우도 있고 산모양에 굴곡등이 포함되는 경우가 많다. 그러나 많은 경우에 이와 비슷한 분포를 가지고 있음이 알려져 있다.

이러한 통계적 특성을 고려하여 SAD 값 연산의 횟수를 대폭 줄인 기법으로 3 단계 검색(Three Step Algorithm)이 있다. 이 알고리즘은 최초에는 저해상도에서 검색을 하다가 차차로 고해상도 검색을 해나가는 방법으로 앞서의 통계적 특성이 일치하는 경우 정확한 움직임 벡터를 찾으면서도 SAD 계산 횟수는 같은 검색 영역을 가지는 전역 검색에 비해 90% 정도의 감축을 가져오게 된다. 그러나 그러한 통계적 특성과 다소 다른 특성을 가지는 검색 영역에 대해서는 Best-match 점과 조금 떨어진 점을 찾을 수 있으나, 비슷한 원리의 다른 알고리즘들에 비해 우수한 성능을 나타낸다.

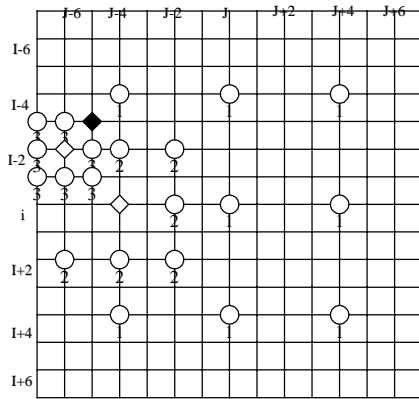


그림 7 3단계 탐색 알고리즘

3 단계 움직임 탐색 알고리즘은 효과적으로 움직임 검색 과정의 계산량을 감소시키도록 해주지만 화상 회의 영상 데이터의 특성을 고려하지는 않았으므로 개선의 여지가 있음을 알 수 있다. 3.1절의 실험 결과에서도 알 수 있듯이 화상 회의 영상 데이터의 주목되는 특징은 주변의 매크로 블록으로부터 예측된 움직임 벡터와 실제 움직임 벡터가 일치하는 비율이 아주 높다는 것이다. 따라서 이러한 예측 벡터와 실제 벡터가 일치하는 매크로 블록을 적은 양의 계산에 의해 찾을 수 있다면 상당부분의 계산량을 줄일 수 있다. 이처럼 예측된 벡터와 실제 움직임 벡터가 일치하는 매크

로 블록을 탐지하여 3단계 움직임 검색 방법을 개선하기 위해서 다음과 같은 방법을 적용하였다.

- i) 예측된 벡터에 해당하는 SAD(Sum of Absolute Difference)값을 구한다.
- ii) 예측된 벡터 주위의 8개의 픽셀에 대한 SAD 값을 구한다.
- iii) i)에서 구한 SAD 값이 주위의 8개 픽셀에 대한 SAD 값보다 적을 때 이 예측된 벡터가 실제 벡터와 일치하는 것으로 간주한다.
- iv) iii)의 경우가 아닌 경우는 다시 3 단계 탐색 알고리즘을 적용하되, 이미 계산된 부분의 SAD 값은 다시 계산하지 않는다.

이 알고리즘을 적용한 경우에 한 매크로 블록의 움직임 벡터를 검색할 때 나타나는 각각의 경우의 벡터들의 SAD 계산량은 다음과 같이 나타난다.

- i) 예측 벡터와 실제 벡터가 일치한 경우 : 9
  - ii) 예측 벡터와 실제 벡터의 차가 1내인 경우 : 25
  - iii) 예측 벡터와 실제 벡터의 차가 2 이상인 경우 : 33
- iii)의 경우는 3단계 검색 기법보다 더 많은 계산량을 필요로 하지만 이를 3절의 화상 회의 영상 데이터의 통계적 특성을 고려하면 우수한 결과를 가짐을 알 수 있다.

이 알고리즘을 적용한 경우의 계산량을 CLAIRE 영상과 같은 전형적 화상 회의 수직적으로 계산하고, CARPHONE이나 FOREMAN과 같이 다소 이러한 특성을 벗어나 경우에 대해 수직적으로 계산하면 다음과 같은 결과를 볼 수 있다.

- i) CLAIRE 영상의 SAD 계산량 :  $0.96X9 + 0.02X25 + 0.02X33 = 9.8$
- ii) CARPHONE 영상의 SAD 계산량 :  $0.66X9 + 0.17X25 + 0.17X33 = 15.8$
- iii) FOREMAN 영상의 SAD 계산량 :  $0.51X9 + 0.21X25 + 0.28X33 = 19.08$

3 단계 검색 알고리즘의 SAD 계산량이 25로 일정하므로 화상 회의 영상 데이터의 특성을 가장 잘 반영한 CLAIRE 영상 샘플의 경우는 3배 가까운 향상을 보이고 다소 그렇지 않은 특성을 가진 영상 샘플에 대해서도 속도 향상 효과가 있음을 수직적인 계산 결과로도 확인할 수 있다.

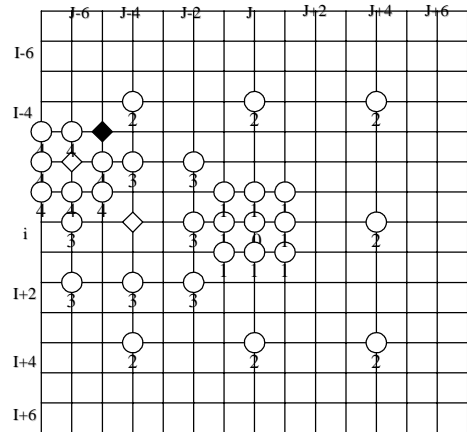


그림 10 개선된 3단계 탐색 기법

그림 11부터 13은 각각의 검색 기법을 적용했을 때 측정된 SAD 함수 호출 횟수와 부호화후 다시 복호화한 영상의 신호대 잡음비(SNR), 그리고 최종 데이터량을 비교한 것이다.

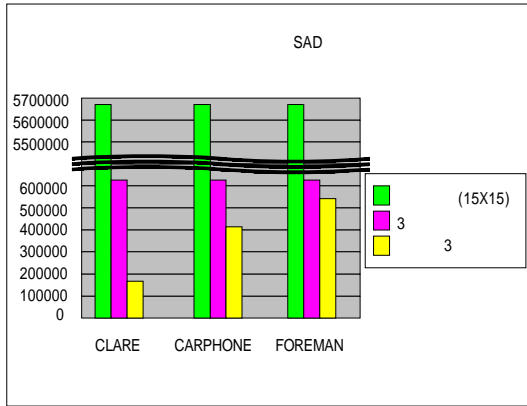


그림 11 검색 알고리즘에 따른 SAD 호출 횟수 비교

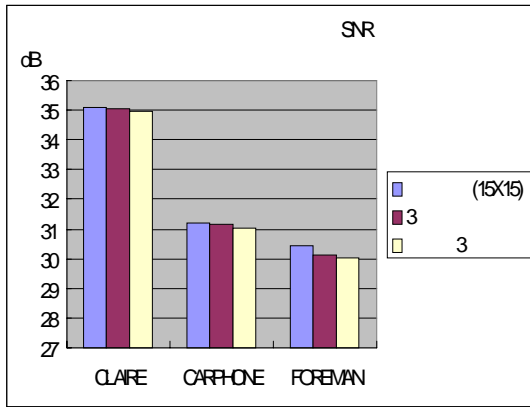


그림 12 검색 알고리즘에 따른 SAD 호출 횟수 비교

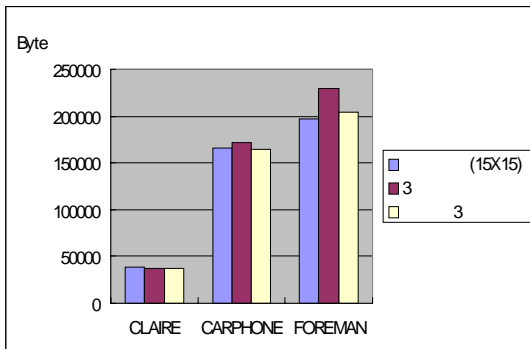


그림13 각 검색 알고리즘에 따른 최종 데이터량 비교

주목되는 점은 개선된 3단계 탐색 알고리즘을 적용한 경우에 그림 11과 같이 화상 회의 성격을 잘 반영한 CLARE 영상에서 3단계 탐색 알고리즘에 비해 처리 속도가 뚜렷이 향상되며, 다소 벗어난 영상들에 대해서도 항상 효과를 보인 것이다. 또한 움직임 벡터 검색의 정확도의 지표인 최종 압축 데이터 량의 경우에도 개선된 3단계 탐색이 3단계 탐색에 비해 진역 탐색의 결과와 더 가까움을 볼 수 있다. 이것은 제안된 개선된 3단계 탐색 기법이 입력 영상의 통계적 특성을 고려하여 확률이 높은 부분의 검색 횟수를 늘렸

기 때문에 더 정확한 움직임 벡터를 찾을 수 있다는 것을 의미한다.

따라서 개선된 3단계 탐색은 화상 회의를 위한 시스템에서 우수한 성능을 발휘할 것이라는 것을 알 수 있다.

### 5. DCT 및 양자화 과정 속도 개선

4절에 기술된 움직임 벡터 검색 개선 알고리즘을 개선하게 되면 움직임 벡터 검색의 계산 과정이 대폭 줄어들어 효과를 가져온다. 그렇게 되면 다른 연산 과정인 DCT나 양자화 계산 과정의 계산량이 큰 부분을 차지하는 결과를 가져오게 된다.

따라서 이러한 DCT나 양자화 과정에서도 개선의 필요성이 요구된다. 여기서 사용될 수 있는 기법은 DCT 연산의 결과로 나타나는 DC 값의 특성을 이용하여 부호화에 사용되지 않을 부분을 검지하여 이 부분에 필요한 DCT 및 양자화 과정을 생략하는 것이다. 즉 DC 값이 다른 AC 값들에 비해 크다는 특성을 활용하여 양자화 처리후의 DC 값이 0보다 작은 경우 다른 AC 값을 구하지 않고 해당 블록의 연산 결과가 모두 0이라는 예측을 적용하는 것이다.

이 DC 값의 계산을 살펴보면,

$$F(0,0) = \frac{1}{8} \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \quad \text{식 1}$$

와 같다. 여기서 양자화 과정에서 사용되는 계수(2Q)를 도입하여 DC 값이 0이 되는 조건, 즉 해당하는 블록 전체가 부호화되지 않을 조건을 구하면 다음 식과 같게 된다.

$$\left| \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \right| < 16Q \quad \text{식 2}$$

이 식을 만족하는 경우의 블록에 대해서는 해당 블록 전체가 부호화되지 않는 것으로 가정하고 해당 블록의 DCT와 양자화 과정을 생략할 수 있다. 그러나 좀더 안전한 가정으로서 다음의 식을 만족하는 경우에 부호화 되지 않도록 할 수 있다.

$$\left| \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \right| < 8Q \quad \text{식 3}$$

이 조건식의 오른쪽 항은 DCT 계산 과정에 이미 포함된 과정이므로 추가적인 계산 과정 없이 조건식만을 추가하여 조건이 만족되면 이후의 DCT 과정이나 그에 따른 양자화 과정을 생략할 수 있게 된다.

### 6. 속도 개선 효과 실험 결과

제안된 기법들을 펜티엄 233MHz급 PC에서 300 프레임의 벤치마크 영상에 대해 적용한 결과, 그림 14와 같이 처리 속도 향상이 2배에서 3배정도에 이르렀다.

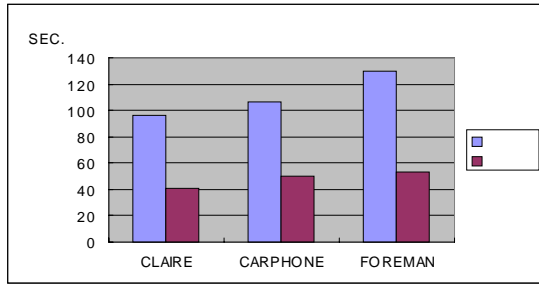


그림 14 처리 속도 개선 효과

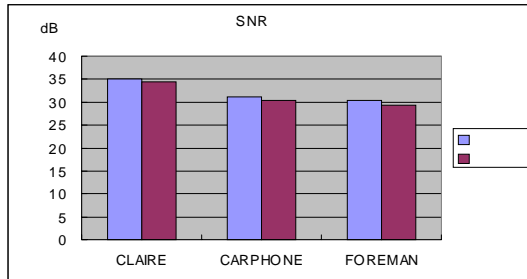


그림 15 SNR 비교

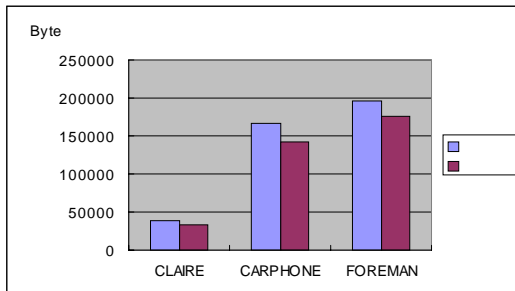


그림 16 최종 압축 데이터량의 비교

또한 최종 압축 데이터량은 오히려 개선전보다 감소하는 결과를 보였다. 이러한 데이터량의 감소는 그림 15에서와 같이 SNR의 감소로서 다소 화질이 저하되는 효과를 보였지만, 실제 화상을 복호화한후 눈으로 확인할 때에는 거의 구별할수 없을 정도의 차이만을 보였다.

## 7. 결론 및 추후 과제

본 연구에서는 인터넷 화상회의를 위한 시스템 구축을 위해 H.263 영상 부호화기를 적용할 때 문제가 될 수 있는 처리 속도에 대한 개선 기법들을 제안하고, 이를 실험을 통해 검증하였다.

제안된 움직임 검색 기법은 화상 회의 영상 입력 데이터의 통계적 특성들을 고려하여 화질이나 압축량에 미치는 영향을 최소화하면서 제거 가능한 연산 과정을 적용적으로 감축함으로써 기존의 3단계 검색 기법보다 우수한 결과를 보였다.

또한 네트워크 상태에 따라 적용적으로 변하는 양자화 계수를 활용하여 움직임 검색의 조기 중단 기법을 도입하여 변환 및 양자화에 걸리는 시간도 단축하였다. 이러한 기법은 다소 화질의 저하는 가져오지만, 눈으로 보기에 구별이 어려울 정도로 미미한 정도이므로 실제 시스템 응용상에 문제가 되지는 않을 것이다.

개선된 알고리즘을 적용한 H.263 부호화기는 기존에 구현된 Telnor 부호화기에 비해 2배에서 3배 정도의 속도 개선

효과를 보였다.

이러한 기법들의 도입의 결과 움직임 예측이나 변환, 양자화 등 핵심적 연산들의 비중은 현저히 감소했으나 다른 요소들의 처리 시간이 상대적으로 증가하는 결과를 가져왔다. 이러한 부분들에 포함되는 것은 검색 영역의 처리를 위한 데이터 이동 루틴이나 복호화된 영상의 재 부호화 과정등이 포함된다. 개인용 컴퓨터도 다중 작업이 기본적으로 지원되고 있는 추세이므로 처리 속도는 더욱 빨라질수록 원활한 작업 환경의 제공이 가능하므로 추후 과제로서 이러한 부분의 처리 속도 개선이 요구된다. 이러한 과정은 적용되는 플랫폼의 하드웨어적인 특성을 활용한 어셈블리의 적용등이 적합할 것이다. 또한 움직임 예측이나 변환 및 양자화 과정에도 어셈블리를 적용하여 더욱 성능을 개선하면 좀더 다양한 환경에서 자연스러운 화상 회의가 가능해질 것이다.

## 8. 참고 문헌

- [1] ITU-T Recommendation H.263 (1996) "VIDEO CODING FOR LOW BITRATE COMMUNICATION"
- [2] A. Yu, R. Lee, and M.Flynn, "Performance Enhancement of H.263 Encoder Based On Zero Coefficient Prediction" ACM MULTIMEDIA 97 The Fifth ACM International Multimedia Conference
- [3] T. Koga et al., "Motion compensated interframe coding for video conferencing," NTC '81, National Telecommun. Conf., pp. G5.3.1-G5.3.5, New Orleans, LA, Nov.-Dec. 1981.
- [4] R.Steinmetz, "Data Compression in Multimedia Computing - Standard and Systems," *Multimedia Systems*, Vol. 1, 1994, pp.187-204
- [5] R. Arrividd et. al., "Image and Video Coding Standards," *AT&T Technical Journal*, Jan./Feb.1993, pp.67-88