

論文2000-37SP-5-9

## 동영역 분할에 의한 고속 움직임 추정 부호화에 관한 연구

## (A Study on the Fast Motion Estimation Coding by Moving Region Segmentation)

李奉鎬\*, 崔敬洙\*, 郭魯尹\*\*, 黃秉元\*\*\*

(Bong-Ho Lee, Kyung-Soo Choi, No-Youn Kwak, and Byong-Won Hwang)

## 요 약

본 논문은 영역분할 정보를 이용한 움직임 추정 기법을 제안하고자 한다. 움직임 추정은 연산량이 매우 방대해서 소프트웨어적인 실시간 구현이 매우 어렵고 전용 칩을 사용하여 구현하고 있다. 본 논문에서는 영역 분할 정보 및 가변 탐색 창 설정을 통해서 FSMA에 비해 연산량을 줄이는 영역 기반 움직임 추정 기법을 제안한다. 둘째로, 분할된 영역 중 얼굴과 같은 의미를 갖는 영역을 선택적으로 부호화하기 위한 인식 기법을 제안하고자 한다. 의미 영역 인식의 기본 취지는 화상 회의나 화상 전화와 같은 대체로 움직임이 적고 화자 중심으로 이루어진 영상에서 화자의 살색 영역 즉, 얼굴 영역을 우선 보호하여 주관적인 화질을 개선하는 데에 있다.

## Abstract

This paper presents motion estimation method using region segmentation information. Motion estimation which is very difficult to be implemented only by software because of intensive computation cost, is implemented by special-purpose hardware in real-time applications. In this paper, we propose region based motion estimation algorithm which can reduce the computation cost by using region segmentation information and setting the variable search window compared with FSMA algorithm. Secondly, another proposed algorithm is to segment semantic region like face for selective coding and transfer of semantic region using segmented region information. This work aims to improving the subjective quality of skin color region or face region in the picture that has slow motion and is mainly composed of one or two speakers of video conference and video telephony applications.

\* 正會員, 韓國電子通信研究院

(Electronics and Telecommunications Research Institute)

\*\* 正會員, 天安大學校 情報通信學部

(Division of Information and Communication Eng., Chonan Univ.)

\*\*\* 正會員, 韓國航空大學校 電子·情報通信·컴퓨터工學部

(School of Electronics, Telecommunications, and Computer Eng., Hankuk Aviation Univ.)

接受日字: 1999年6月1日, 수정완료일: 2000年4月28日

## I. 서 론

디지털 이동 통신이나 PSTN을 이용한 비디오 데이터의 전송은 매우 제한적인 대역을 갖음으로 인해서 64kbps이하의 데이터 압축이 요구되는 실정이며 이를 위한 많은 영상 압축 기법이 연구 중이다. 현행 ITU-T의 H.26x 계열의 압축 표준안 및 MPEG-4 시스템에서는 DCT를 기반으로 이루어지는 상태이고 보다 나은 부호화 효율을 위해서 영상에 의미를 부여하여 부호화

하는 의미 기반 부호화(semantic-based coding), 영역 기반 부호화(region-based coding), 특정한 모델 설정을 통한 모델 기반 부호화(model-based coding), 객체기반 부호화(object-based coding) 등이 연구되어 지고 있는 상태이다. MPEG-4에서는 MPEG-2에서와 같은 프레임 단위가 아닌 객체 추출을 통한 VOP(Video Object Plane)별로 부호화한다. 본 논문에서는 H.261을 기본으로 분할기반 부호화(segment-based coding)기법을 적용해 성능 개선을 보이고자 한다.

우선, 몇 가지 개선해야될 문제점을 살펴보면, 움직임 추정은 부호화 성능에 있어서 매우 중요한 기법으로서 움직임 추정이 잘 이루어지면 이루어질수록 에러 이미지의 부호화 비트율이 낮아지므로 H.26x에서는 움직임 추정을 위한 탐색 방법으로 전역 탐색 기법(FSBMA: Full Search Block Matching Algorithm) 및 TSS(Three Step Search) 기법 사용을 주로 권고하고 있지만 이들 기법은 탐색 시 연산량이 많은 관계로 실시간 응용 시 많은 어려움이 있다.<sup>[1]</sup> 이를 위해 본 논문에서는 영역 분할 정보를 이용해 구성된 영역별로 탐색 영역을 가변적으로 적용해 연산량을 줄이는 분할 기반 움직임 추정(segment-based motion estimation)기법을 제안하고 있다. 다음으로는 초 저속용 비디오 부호화는 많은 가변 분해능 및 적절한 비트율 제어 알고리즘이 요구되는데 H.26x 기반으로 분할된 영역의 분석을 통한 얼굴이나 손과 같은 의미 영역에 대한 선택적인 부호화를 지원하며 이를 선별적으로 전송하는데 쓰일 수 있는 인식 기법을 제안하고자 한다. 본 논문의 구성은 다음과 같다. II장에서는 기존의 움직임 추정 기법들에 대해서 살펴보고, III장의 3.1절에서는 분할 기반 움직임 추정(segmented-based motion estimation) 기법을 기술하며, 3.2절에서는 B-Picture의 영역별 움직임 정보 부호화를 제안하고, 3.3절에서는 의미를 갖는 영역의 우선적 전송을 위한 인식 기법을 제안하며, IV장에서는 실험 결과 및 검토 과정을 거치며, 마지막으로, V장에서는 결론에 대해 기술한다.

## II. 기존의 움직임 추정 기법

움직임 보상 방식의 핵심 기술은 움직임 블록을 기반으로 하는 블록 정합 벡터를 검출하는 것으로 일반적으로 블록 정합 방식(BMA : Block Matching

Algorithm)과 화소를 기반으로 하는 화소 순환 방식(PRA: Pel Recursive Algorithm), 위상 상관 방식, 시공간 경사 방식(gradient based method) 등으로 구분된다. 영상 내에서의 패닝, 축소/확대, 회전과 같은 복잡한 움직임에 대한 검출의 정확도 면에서는 PRA 방식이 BMA를 성능 면에서 능가하나 연산량 및 시스템 구성 시의 복잡도 측면에서는  $16 \times 16$ 이나  $8 \times 8$  블록단위로 처리하는 BMA가 구현이 용이하다. 이에 따라 H.26x 및 MPEG-1, MPEG-2, MPEG-4에서는 PRA보다 예측 오차 성능이 떨어지지만 하드웨어 구성이 용이한 BMA 방식을 움직임 추정 부호화 기법으로 채택하고 있다. 블록 정합 알고리즘은 한 블록내의 모든 화소들이 같은 방향으로 평행으로 이동한다는 것을 가정하여 움직임 추정을 하는 방법으로서, 한 영상을 여러 개의 작은 블록으로 분할한 후, 움직임 탐색 기법을 이용하여 각 블록 단위로 움직임 벡터를 찾아 그 벡터를 해당 블록의 화소 전체에 적용하는 방법이다. 즉, 동일 블록 내에 존재하는 모든 화소는 동일한 움직임을 갖는 강체(rigid body)라는 가정이 필요한데, 이 가정은 BMA의 성능 면에서 단점이자 장점이 되는 특성이다. BMA의 움직임 추정 기법을 살펴보면, 움직임 벡터를 찾기 위하여 먼저 현재 프레임의  $(N \times N)$  크기의 블록으로 겹치지 않게 나누고 참조 프레임의  $(2w+N) \times (2w+N)$  크기의 탐색 영역내에서 후보 블록들과 비교한다. 현재 블록과 후보 블록들과 비교 후 왜곡 함수 또는 정합 기준의 최소 오차를 갖는 후보 블록과 현재 블록과의 변위 차이를 해당 블록의 움직임 벡터로 설정한다. 일반적으로 움직임 벡터를 찾기 위해서 전역탐색 기법이 많이 사용되고 있는데 전역 탐색에서는 움직임 벡터를 얻기 위해서는  $2(2w+1)$ 의 후보 블록들과 정합을 하여야 한다. 현 블록과 예측 후보 블록 사이의 부호화를 측정하는 비용함수는 여러 가지가 이용될 수 있으나 일반적으로 평균 자승 오차(MSE; Mean Square Error), 평균 절대 오차(MAD; Mean Absolute Difference), NCCF(Normalized Cross-Correlation Function)가 많이 사용된다. BMA의 기본이 되는 기법은 전역 탐색 블록 정합 알고리즘(FSMA; Full Search Block Matching Algorithm)으로 이는 정합 기준 함수의 최소 왜곡을 갖는 블록의 변위를 찾기 위한 많은 연산량 때문에 실시간 비디오 부호화기를 개발하는데 많은 어려움이 있으며, 대부분 하드웨어로 구현을 하고 있다. 전역 탐색 방법의 연산량을 줄이는 방법은 크게

3가지로 분류할 수 있다.

첫째로, 블록 내에서 정합에 참여하는 화소의 수를 줄이는 방법으로 이는 특정 패턴을 이용하거나 다운 샘플링을 하여 블록의 크기를 줄여 하나의 블록에서 연산량을 줄이는 방법이다. 그러나 정합에 사용되는 화소들이 블록의 특성을 잘 반영할 경우는 좋은 성능을 얻을 수 있지만 그렇지 못한 경우 심한 왜곡이 발생한다. 두 번째로, FSMA의 연산량을 줄이기 위한 연구로, 현재까지 가장 많은 연구가 된 부분으로 탐색 영역 내에서 정합에 참여하는 후보 블록들을 줄이는 방법으로 TSS(Three Step Search), NTSS(New Three Step Search), 2-D LOG(2-D Logarithm Search), OTS(One at a Time Search Algorithm), 1-D FS(1-D Full Search), CSA(Cross Search Algorithm) 등과 같은 많은 빠른 탐색 기법들이 제안되었다. 현재도 FSMA와 유사한 왜곡을 가지며, 연산량을 줄이기 위한 연구가 활발히 진행 중에 있으나, 빠른 탐색 기법들은 최소 왜곡을 갖는 곳으로부터 멀어질수록 왜곡이 증가한다는 가정 하에서 개발되었기 때문에 국부 최소값(local minima)에 수렴하는 문제점이 존재한다. 따라서, 이러한 고속 블록 정합 알고리즘들은 전역 탐색에 비해 추정 오차가 크고, 추정된 움직임 벡터의 신뢰성이 현저히 떨어지는 문제점이 있다. 이에 따라 추정된 움직임 벡터에 대한 신뢰성을 최대한 유지하면서 탐색 영역과 연산량에 대한 문제를 해결하기 위해 계층적 탐색 방식이 제안되었지만, 이 방식 또한 기존의 고속 알고리즘에서와 같이 서로 다른 해상도를 갖는 계층들 간에서 점진적으로 움직임을 추정해 가는 방식이기 때문에 움직임 추정 벡터가 지역적 최소값(local minima)에 빠질 가능성을 배제할 수 없다. 국부 최소값에 수렴하는 경우 움직임 보상 후 예측 오차가 상대적으로 커지기 때문에 오차를 부호화 할 경우 많은 비트 자원을 사용하여야 한다. 마지막으로, 다른 방법으로는 연산량을 줄이기 위해서 탐색 영역의 크기를 동적으로 조정하는 기법들이 다수 제안되어 있다<sup>[11]</sup>. 추정 블록의 크기가 고정된 기존의 블록 정합 알고리즘의 단점을 극복하기 위해, 국부적인 움직임의 분포 양상에 따라 움직임 추정에 사용하는 블록의 크기를 가변적으로 설정함으로써 움직임 추정의 예측 성능을 향상시킬 수 있다는 사실이 다수 보고되고 있다. 이에 기반하여 블록 크기를 가변시키는 가변 블록 정합 알고리즘이 다양한 접근 방법에 의해 연구되어 왔으나, 블록 크기를 크게 하여

움직임을 추정한 후, 예측 오차가 크면 블록을 분할하는 기존의 방식은 중복된 계산으로 인해 처리 시간 측면에서 효율적이지 못하고, 또한, 블록 크기가 너무 다양할 경우에는 부호화 시에 복잡도가 증가하게 되고 전송 부담이 증가하는 등의 문제가 있다<sup>[12-13]</sup>.

### III. 제안된 움직임 추정 부호화

#### 1. 분할기반 움직임 추정

기존의 움직임 추정 기법은 각 매크로 블록별로  $\pm 16$  화소 정도의 탐색 영역을 설정해 움직임을 추정하는 것이 가장 일반적이다. 그러나 화상 회의나 화상 전화 응용의 경우에 있어서는 통계적으로 작은 움직임을 갖는 특성이 있어  $\pm 16$ 범위로 움직임을 추정하는 데에는 비효율적인 면이 있고 계산상에도 많은 계산 시간이 필요하다<sup>[2-3]</sup>. 고속 알고리즘으로 3단계 탐색 기법(TSS) 및 계층적 움직임 추정 기법(HBMA) 등이 연구되었지만 본 논문에서는 영역을 분할하는 기법을 적용하여 움직임을 갖는 영역들을 대략적으로 분할한 후, 이 분할 정보를 이용해 분할된 영역 중 임의로 몇 블록을 추출해  $\pm 16$ 까지 움직임을 추정해 그 영역이 갖는 움직임 범위를 가변적으로 설정하여 고속 탐색을 실현하였다. 여기에서의 움직임 추정은 역방향(backward)에만 국한되며 양방향 예측을 위한 순방향(forward) 움직임 예측은 고려하지 않는다. 움직임 추정과정은 다음과 같다.

① 가우시안 필터링 : 두 영상에 대해서 움직임의 상관을 개선하기 위해  $3 \times 3$  저주파 가우시안 필터링을 취한다. 이 과정은 움직임 추정시 영상이 가지고 있는 불규칙 잡음을 없애줌으로써 움직임 벡터 추정 성능을 다소 개선할 수 있는 장점이 있지만 실시간 적용시 다소의 계산상의 지연이 발생하는 부분으로서 계산상의 지연이 문제가 되는 경우에는 적용하지 않아도 무방하다.

② 두 프레임 간의 차영상을 식(1)과 같이 구한다.

$$\begin{aligned} \text{DiffPixel}(x, y) &= |I_k(x, y) - I_{k-1}(x, y)| \\ \text{where, } I_k(x, y), & \text{ Current Image,} \\ I_{k-1}(x, y), & \text{ Previous Image} \end{aligned} \quad (1)$$

③ 영역 병합 기법을 이용한 영역 분할 : 차영상에 대해 8×8 블록으로 나눈 후, 임계치(Th1)이하의 차 화소수를 가진 블록은 제거하고 영역 병합 기법을 이용해 움직임을 갖는 영역을 구성한다. 병합 시 8 방향을 고려한다.<sup>[4]</sup> 이 과정에서 마커(marker)를 추출한다.

```
if( M(Bi) == TRUE)
RegionMerging( );
MarkerExtraction( );
```

여기서,  $M(B_i)$ 는 블록  $i$ 의 병합 조건이다. 병합시 차 영상에 나타나는 고립된 작은 블록들은 잡음에 의한 블록에 가깝고 그다지 영상에서 의미를 갖지 않으므로 제거한다.

④ 구성된 영역들에 관한 레이블 정보를 추출한다.

⑤ 영역별 움직임 탐색 영역의 설정 : 먼저 구성된 영역들에 대해서 영역을 포함하는 최소의 사각형 좌표를 구한다. 이 사각형 위치 정보를 이용해서 탐색 영역을 구하고 움직임 벡터를 부호화 하게 된다.

⑥ 샘플 블록의 추출 및 움직임 추정 : 분할된 영역들의 블록 수에 비례하여 샘플 블록을 임의로 추출한 후, 식 (2)에서와 같이, ±16 범위로 MAD(Mean Absolute Difference)를 이용해 움직임을 추정한다.

$$MAD(i, j) = \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |I_t(x+m, y+n) - I_{t-1}(x+m+i, y+n+j)| \quad (2)$$

많은 샘플 블록을 추출할수록 실제 탐색 영역을 찾을 가능성은 커지나 고속 추정을 목표로 하고 있으므로 적당한 수준에서 샘플을 추출한다.

⑦ 가변 탐색 영역 결정 : 영역별로 추정된 움직임 벡터들 중 가장 값이 큰 벡터들부터 신뢰도를 측정한 후 가변 탐색 영역을 설정한다. 신뢰도 측정 과정은 구성된 블록들의 평균 에러 및 분산 정보를 이용해서 결정한다. 이 과정에서 움직임 영역으로 구성되지 않은 블록들은 작은 움직임을 갖는다는 가정 하에 탐색 영역을 ±2까지 설정해서 움직임을 추정하고 보다 연산량을 낮추기 위해서는 추정하지 않아도 무방하다.

⑧ 고속 움직임 추정 : 각 영역별로 구성된 가변 탐색 영역을 이용해 8×8 블록별로 설정된 탐색 영역 범

위 내에서 움직임 추정을 수행한다. 이 과정에서 앞의 단계에서 탐색한 블록들은 중복 탐색하지 않는다. 움직임 추정을 위한 비용 함수는 MAD를 이용한다.

⑨ 확장 재 추정 : 추정된 블록 중 에러가 임계치(Th2)보다 크면 ±16까지 확장하여 재 추정한다.

```
if(MadError( Bi ) > Th2)
ExpandEstimate( );
```

확장 추정 탐색 범위는 그림 1에 예시되어 있으며 확장하여 재 추정을 하는 목적은 가변 탐색 영역 설정이 잘못되었을 경우 발생하는 국부 최소를 극복하기 위한 것으로서 이 과정에서 연산량이 다소 늘어나며 움직임이 큰 영상일수록 이 과정에서의 연산량이 늘어나게 된다.

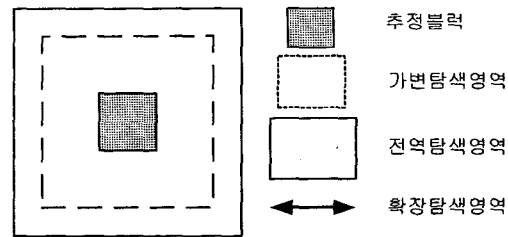


그림 1. 확장 탐색 영역  
Fig. 1. Extended search area.

2. 영역 재병합 및 움직임 벡터 부호화

H.263은 IBP의 픽취 구조를 갖는다. I 프레임은 Intra로 부호화되고 P는 Intra와 Inter로 B는 Inter 부호화가 권고되는데, 이 절에서는 B 프레임 부호화 시 성능 향상을 위해 분할된 영역 정보를 이용해 B-Picture의 움직임 정보를 부호화 하는 단계로 영역 재 병합 과정과 최종의 마커 추출 및 영역의 위치 정보 추출 후 허프만 부호화를 통해 움직임 정보를 부호화하게 된다. 영역 재 병합(region re-merging)은 움직임 추정을 위해서 전 단계에서 구성된 분할 영역에 추정된 움직임 벡터를 이용해 재구성하게 된다. 재 병합 과정은 다음과 같다.

① 움직임 정보를 이용한 병합 : 움직임 추정을 위해서 전 단계에서 구성된 분할 영역에 추정된 움직임 벡터를 이용해 재구성하게 된다. 즉, 병합되지 않은 블록

중 제로가 아닌 벡터를 갖는 블록은 전에 구성된 영역과 8방향으로 인접하면 병합하고 구성된 블록의 내부에 속한 블록들도 움직임을 갖는 영역의 일부분으로 간주하여 병합하여 마커를 갱신시킨다.

이 단계에서도 제로 벡터가 아닌 작은 고립 영역들은 3.1의 단계 3에서처럼 병합되지 아니하고 제거된다.

② 구성된 영역의 마커 및 위치 정보 추출 : 재 병합된 영역들이 서로 8방향으로 연결되면 다시 재 병합 과정을 거친다. 마지막으로 마커를 갱신시키고 각 영역 별로 영역을 포함하는 최소의 사각형 위치 정보를 구한다. 재 병합과정은 마커를 갱신하는 관계로 소프트웨어적인 구현은 그다지 계산량에서 그다지 부담으로 작용을 하지 않는다.

③ 움직임 벡터 부호화 : 움직임 벡터의 부호화는 허프만 부호화를 통해 통계적인 압축을 수행하는 것이 일반적인 데, 본 연구에서는 H.263에서 적용하는 예측 코딩(differential coding)기법을 이용한다<sup>[13]</sup>.

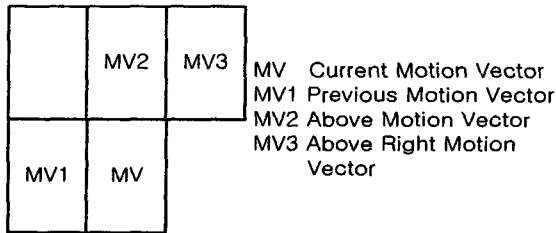


그림 2. 예측 코딩 기법  
 Fig. 2. Differential coding method.

$$\begin{aligned}
 P_x &= \text{Median}(MV1_x, MV2_x, MV3_x), \\
 P_y &= \text{Median}(MV1_y, MV2_y, MV3_y) \\
 MVD_x &= MV_x - P_x, \\
 MVD_y &= MV_y - P_y;
 \end{aligned}$$

④ B-Picture의 움직임 부호화 : 움직임 벡터의 부호화는 프레임별로 부호화 하지 않고 분할된 영역별로 위치 정보 및 움직임 벡터를 부호화 한다. 위치정보는 영역을 포함하는 사각형 정보로서 CIF 이미지인 경우는 각 위치 당 6bit만을 할당하면 총 2BYTE면 표현 가능하다. 신택스 구조는 그림 2와 같다.

H.263 기반의 B-Picture에 비해 비트율이 적어지며, 구성된 영역의 수가 적고 영역의 크기가 적을수록 비

트 효율이 커진다.

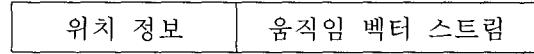


그림 3. 움직임 벡터 신택스 구조  
 Fig. 3. Syntax of motion vector.

⑤ 에러 이미지 보상 부호화 : 움직임 추정 후 생기는 에러 이미지는 기존의 H.26x나 MPEG-1,2,4에서 처럼 DCT기법을 이용해 처리를 한다. 본 논문에서는 분할된 영역들에 속하는 블록들에 대해서만 에러 보상을 수행하는데 에러가 심한 블록만을 보상한다. 기존의 방식과는 다르게 매크로 블록별로 수행하지 않고 다음에 설명할 가변 영역 분해능 부호화(scalable region coding)를 지원하기 위해서 영역별로 부호화 된다.

3. 가변 영역 분해능을 위한 의미 영역 인식기법

화상회의나 화상전화 같은 영상은 크게 전경(background image)과 배경(background Image)으로 구성되는데 전경은 배경에 비해 의미가 크다고 볼 수 있다. 또한 전경을 얼굴과 몸통 부분으로 세분해보면 얼굴 부분이 의미를 더 가진다고 볼 수 있다. 이러한 점에 착안해서 움직임을 갖는 영역들을 분할하고 이 분할된 정보를 이용해 얼굴 영역과 손 영역을 인식해 제한된 채널 환경에서도 주관적인 화질을 유지할 수 있다. 실제로 위에서 제시한 영역 분할 결과를 살펴보면 구성된 영역들은 얼굴 영역과 손을 중심으로 한 신체영역으로 분할된다. 기존의 H.263이나 MPEG-4에서는 다양한 시공간 가변 분해능(spatial and temporal scalability)을 지원하는데, 이 경우 다소 전체영상의 화질 저하를 초래하는 단점이 있다. 그러므로, 분할된 영역중 의미가 큰 얼굴과 손을 중심으로 한 영역들에 우선권을 부여해 개별적인 부호화 및 전송을 위한 인식 기법을 제시한다. 인식은 살색 정보를 이용하며 색 정보를 이용한 인식은 영상 분석에서 종종 쓰이는 기법으로 MPEG-4가 요구하는 VOP 인식을 위해서도 많이 쓰이고 있으며 객체 기반 부호화에서도 얼굴 영역을 인식하기 위해서 종종 사용되고 있다.<sup>[6]</sup> 제안하고자 하는 살색 영역 인식 및 우선권 부여 과정은 그림 3과 같다. 세부적인 알고리즘은 다음과 같다.

① 색 신호 성분 U, V를 이용한 살색 영역 인식 :

사람의 피부색은 인종마다 다소의 차이가 있지만 색 신호 좌표 계에서 보면 일정한 영역에 분포하게 된다.<sup>[5]</sup> 이 영역을 이용해 분할된 영역들에 대해서 블록 내부에 속하는 화소들을 살색화소로 인식한다. 살색 좌표 계는 아래 그림 4와 같다. 그림 4에서의 타원 영역은 색 신호가 존재하는 영역으로 살색화소 결정 여부는 식 (3)을 만족하면 살색 화소로 간주한다.

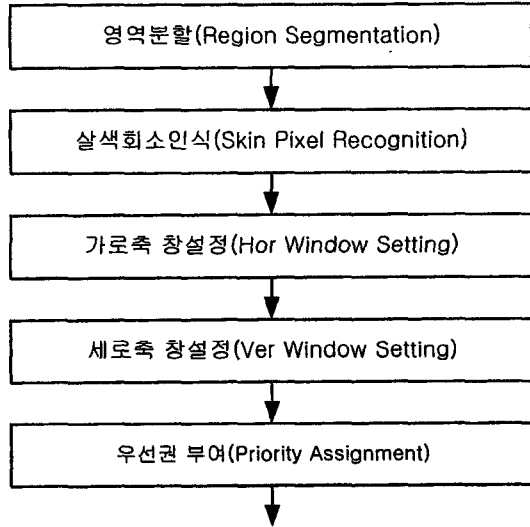


그림 4. 의미영역 인식  
Fig. 4. Semantic area recognition.

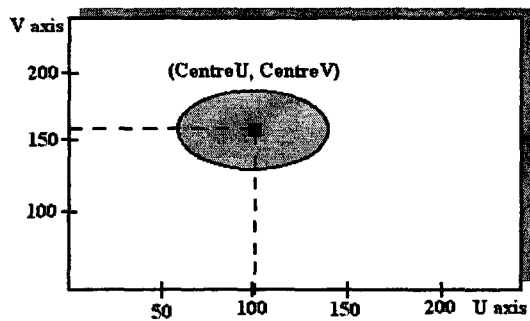


그림 5. 살색 좌표계  
Fig. 5. Coordinate system of skin color.

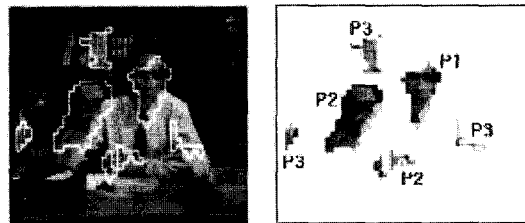
$$\frac{(U - CentreU)^2}{RadU^2} + \frac{(V - CentreV)^2}{RadV^2} < 1 \quad (3)$$

여기서,  $U$  및  $V$ 는 화소의  $U$ 값 및  $V$ 값이고,  $CentreU$  및  $CentreV$ 는 타원의 중심 좌표값이며,  $RadU$  및  $RadV$ 는 타원의 축 반지름을 의미한다.

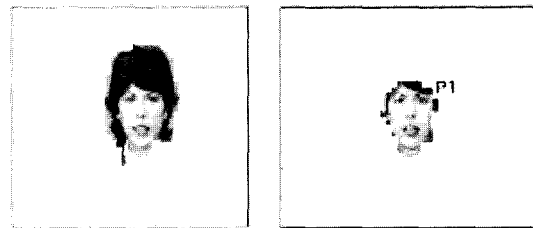
다음으로, 살색 블록 결정 여부는 블록 내부의 화소 수가 미리 정한 임계치(Th3) 이상일 경우에 살색 블록으로 결정하며 마지막으로, 살색 영역은 각 영역의 살색 블록을 하나 이상 가지고 있으면 살색 영역으로 결정한다.

② 가로축 창(Window)설정을 통한 대략적인 얼굴 영역의 분석 : 얼굴 영역은 일반적으로 화면 중앙에 위치하므로 가로축으로 창을 설정해 단계 1에서 찾은 살색 영역 중 창 내부에 속한 분할 영역(segmented region)들을 결정한다.

③ 세로축 창 설정을 통한 얼굴영역 결정 : 위 단계들을 통해서 판정된 살색 영역들이 얼굴이 아닌 손 같은 다른 영역일 수 있으므로 세로축으로 창을 설정해 얼굴 영역을 결정한다.



(a) salesman



(b) claire

그림 6. 우선권을 갖는 의미 영역  
Fig. 6. Semantic area with priority.

④ 이렇게 결정된 얼굴 영역에 대해서는 최우선적으로 부호화하고 전송 시 우선권을 주고 나머지 손과 같은 살색 영역에 대해서는 배경과 같은 의미가 덜한 영역에 비해 우선권을 부여한다. 그림 5는 우선권을 갖는 의미 영역을 Salesman과 Claire 이미지에 대해 예시한 것이며 아래 식은 결정된 각 영역에 대해서 우선권을 부여하기 위한 식이다.

Face Region → High\_Priority  
 Non Face Skin Region → Middle\_Priority  
 Non Skin Region → Low\_Priority (4)

#### IV. 시뮬레이션 결과 및 고찰

제안한 알고리즘의 성능 평가를 위해 352×288 Salesman(30Frame), 360×288 Claire(80Frame), 360×288 Miss America(40 Frame) 실험 이미지를 사용하였으며 연속하는 두 개의 영상에 대해서 시뮬레이션을 수행하였다. 제안한 알고리즘 중 분할 기반 움직임 추정의 결과 측정은 전역 탐색 기법(FSMA)과의 성능 평가 지수인 매칭되는 블록 수와 PSNR로 표 1에 나타내고 이에 대한 결과를 그림 8, 9에 예시 하였다. 가변 영역 결정 시 이용한 영역 분할 영상 및 결과는 Salesman 영상에 대해서 그림 6에 예시되어 있다. 그림 8,9에서 볼 수 있듯이 Claire 영상의 경우 제안한 알고리즘이 전역 탐색 기법에 비해 평균적으로 블록 매칭 수가 10.3배 작게 나타났고, PSNR은 제안한 알고리즘이 0.4dB 정도 낮게 나타났다. 전체 결과는 표 1에서 볼 수 있듯이 매칭된 블록 수는 제안한 알고리즘이 작게 나타났고 PSNR면에서는 0.2~0.6dB 정도 낮게 나타났다. 하지만 움직임 추정 속도를 보면 전역 탐색 방법에 비해 탐색 속도는 빠르게 개선할 수 있었지만, 분할 영역이 크거나 움직임이 크고 또는 탐색 범위가 크게 잡힐 경우에는 추정 속도가 일정하지 않고 프레임마다 다르게 발생함을 확인할 수 있었다. 또한 움직임 추정 성능은 영역 분할 결과에 따라 상당히 좌우가 되는데 컴퓨터 시뮬레이션에 사용된 영상들은 대부분 움직임이 적으며 이러한 영상들은 전경(사람)을 중심으로 실제 움직임이 있는 영역들만 분할이 효율적으로 이루어지거나 패닝(panning)이나 일종의 손 떨림이 있는 영상에 있어서는 분할이 프레임의 전 영역에 걸쳐 이루어지므로 다소 효율이 떨어짐을 확인할 수 있었다.

다음으로 제안한 알고리즘인 B-Picture 움직임 부호화 결과는 그림 10 및 표 2에 예시되어 있다. 시뮬레이션은 H.263의 전 실패스를 구현하지 아니하고 발생하는 움직임 정보에 대한 발생 비트량과 제안한 방법에 의해서 발생하는 비트량을 비교한 수치로서 표2에서 볼 수 있듯이 제안한 알고리즘이 FSMA에 비해서 움

표 1. 평균 블록 매칭수 및 PSNR  
 Table 1. Average number of matching block and PSNR.

시험 영상	평균 블록 매칭 수		평균 PSNR	
	FSMA	Proposed	FSMA	Proposed
Claire	1448272	111029	47.77	47.52
Miss America	1448272	111746	44.25	43.65
Salesman	1415040	302452	43.01	42.94

표 2. 움직임 정보 부호화 비트량  
 Table 2. The number of bits for motion information coding.

시험 영상	평균 부호화 비트량(BYTE)	
	FSMA	Proposed
Claire	1209.9	701.48
Miss America	1532	684
Salesman	990	883.31

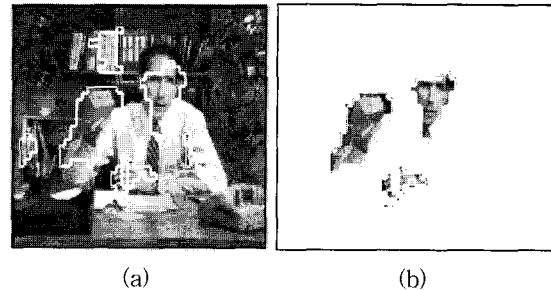
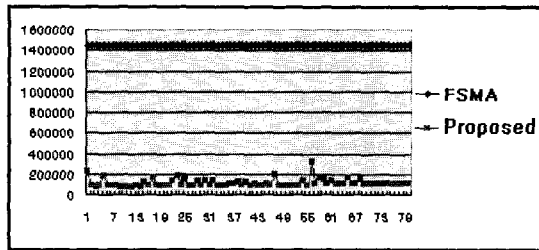


그림 7. 살색 영역 인식  
 Fig. 7. Skin color recognition.

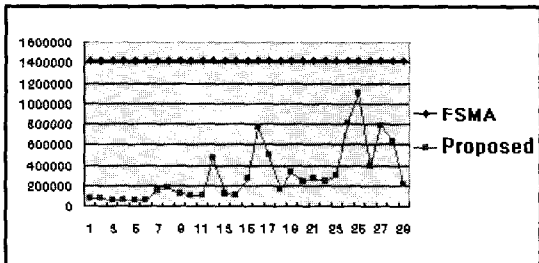
직임 부호 발생 비트량이 적게 나타났다. 움직임 벡터만을 전송하는 B 화면에만 적용한다는 가정 하에서 시뮬레이션의 구체적인 결과를 살펴보면 Salesman 영상의 경우에는 프레임 중간에 손상된 실험 영상에 의한 구성된 영역이 복잡해져 이 부분에서 비트량이 많이 발생해 압축이 크게 되지 않는 결과를 나타냈다. 다음으로 Claire는 약 1.8배 정도의 발생 비트량이 줄어드는 결과를 얻을 수가 있었으며 반대로 PSNR은 0.25dB로 성능이 저하됨을 확인할 수 있었으나 그다지 부호화 압축 성능에 대비 저하가 적으며 Miss America 영상의 경우에도 압축 효과는 2.1배 향상을 가져왔고

PSNR은 0.45dB 저하된 결과를 볼 수가 있었다.

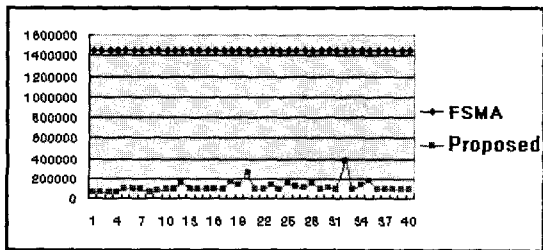
마지막으로 의미를 갖는 얼굴 및 손 영역 인식의 결과영상은 그림 7에 예시되어 있다. 모의 실험은 분할 영역 중 얼굴 영역을 먼저 인식하고 얼굴 영역에 대해서 최우선권을 부여하고 다음으로 결정된 손 영역에 대해서는 다음 우선권을 부여했고 이를 이용해서 선택적으로 부호화를 수행했다. 제안한 병합 알고리즘은 화소단위나 작은 블록 단위의 자세한 셀 병합이 아닌 움직임 추정 단위의 블록을 사용한다. 살색 정보를 이용한 얼굴 영역의 인식은 많은 시험 영상에 대해서 적용을 해 본 결과 조명에 의한 오인식이 발생하는 현상을 볼 수가 있었으며 이 부분에 대해서는 적응적인 살색 화소 선정 알고리즘이 보완되어야 한다. 그림 6은 인식된 살색 영역을 도시한 것으로써 병합된 영역 중에서 Salesman의 얼굴과 손만이 효과적으로 인식됨을 확인할 수가 있었다.



Claire's Matching count

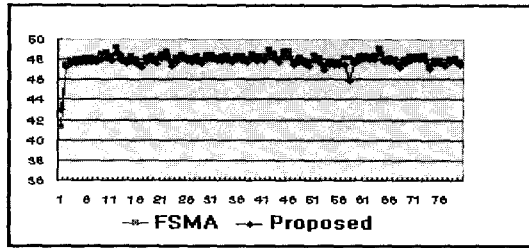


Salesman's Matching count

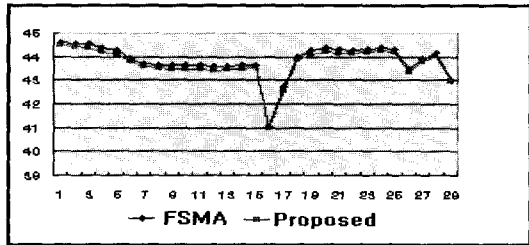


MissAmerica's Matching count

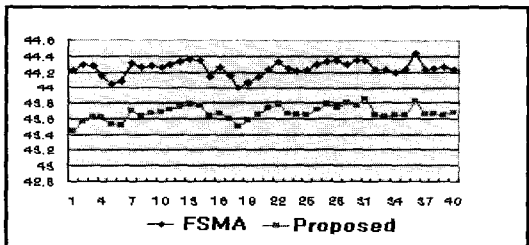
그림 8. 블록 매칭 결과  
Fig. 8. Block Matching Result.



Claire's PSNR After MY Estimation



Salesman's PSNR After MY Estimation



Miss America's PSNR After MY Estimation

그림 9. PSNR 결과  
Fig. 9. Comparison of PSNR Result.

### V. 결 론

본 논문의 알고리즘은 대역폭이 매우 제한된 이동통신 망의 초 저속 영상 부호화에 적용하기 위한 목적으로 제안되었다. 제안된 방법에서는 화상 전화와 같은 영상의 움직임 영역은 비교적 협소한 변위를 갖는다는 사실에 기반을 두어 영역 분할 정보를 이용해 고속으로 움직임 벡터를 추정한다. 그 결과, 제안된 방법은 FSMA에 비해 매우 빠른 탐색 속도를 제공했으며 PSNR은 소폭 감소됨을 확인할 수 있었다. 다음으로, B-Picture 움직임 정보 부호화 알고리즘에 있어서도 분할된 영역 정보를 이용하여 부호화 하는 기법이 기존의 기법보다 비트율이 약 1/2에 정도 적게 나타났다.

전송 대역이 매우 한정된 채널에 영상 정보를 전송할 경우에는 영상의 원 화질 유지보다는 필요한 정보만을 전송하여 부호화 효율을 높여야만 한다. 이에 대한 대표적인 응용으로 MPEG-4의 가상 미팅(virtual



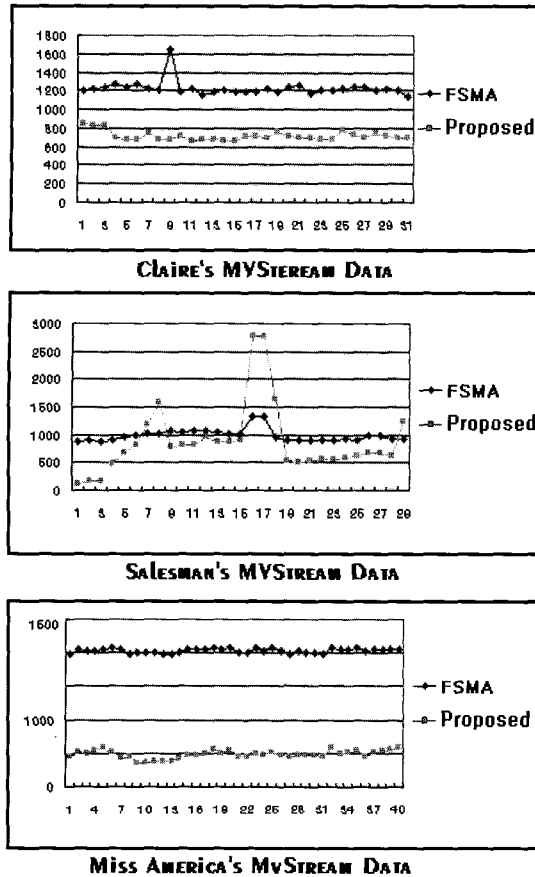


그림 10. 발생 움직임 정보량

Fig. 10. Generated Motion Vector Stream.

meeting) 응용에서는 Facial Animation 기법이 이용되는데 이는 FAP(Facial Animation Parameter)를 이용해 2kbps까지 정보를 압축한다<sup>[2]</sup>. 마찬가지로 제한된 채널 상에서 영상 통신을 할 경우에는 영상 전체의 화질 유지보다는 중요한 부분만을 전송하여 주관적 화질을 유지함으로써 부호화 효율을 높일 수가 있다. 본 논문은 이런 개념을 응용한 것으로 낮은 비트율에서 영상의 전경에 해당하는 얼굴 영역의 화질을 보호해 주관적인 성능을 향상시킬 수가 있었다. 또한 이 기법은 사용자에 의해 우선권이 주어지는 영상 객체나 영상의 일정 영역에 대해서도 적용 가능하다. 하지만 손떨림이나 손실이 있는 영상의 경우는 영역 분할이 복잡하게 이루어져 발생 비트율이 커지는 단점이 있으므로 이 부분의 보상 병합 방법과 같은 보완이 필요하다. 마지막으로, 얼굴이나 손 같은 의미 영역의 인식은 임계치의 대략적인 설정 하에서도 제대로 이루어짐을 확인할 수 있었다. 차후 연구 과제로는 영역별 가변 분해능(region

scalability)을 지원하기 위해 일정한 비트량 생성에 필요한

Rate-Distortion 측면에서의 추가적인 연구가 필요하다. 향후 멀티 미디어가 지향하는 서비스는 사용자의 환경을 보다 친숙하고 강력하게 만들며 다양한 미디어를 지원하고 영상의 경우에는 프로세서의 성능 향상에 따라 고등 알고리즘을 적용하며 영상에 의미를 부여하는 환경으로 진화하고 있다. 영상에 의미를 부여하는 과정은 많은 처리 시간과 인식이라는 복잡한 과정이 필요하나 동일한 환경에서 선택적인 표현이 가능하고 사용자의 편집 기능 및 조작성을 가능케 한다. 본 논문에서 제안한 기법들은 이러한 추세에 다소 근접한 것으로서 향후 실시간 화자 인식, 얼굴 영역 및 의미 영역 인식 기법의 적용이 필요한 시스템의 경우에는 제안한 기법들을 기반으로 응용이 가능하리라 생각한다.

## 참 고 문 헌

- [1] K.R RAO, J.J HWANG "Techniques and Standards for Image, Video, and Audio Coding", *Prentice Hall* 1996.
- [2] Y. Baek, H.-S.Oh, and H.K Lee "Block Matching Criterion for Efficient VLSI Implementation", *Electronic Letters*, 32(13), 1184-1185 (jun.1996)
- [3] B. Liu and A.Zaccarin, "New Fast Algorithms for the Estimation of Block Motion Vectors" *IEEE Trans. On Circuits and Systems for Video Technology* 3(2), 148-157, 9Apr.1993)
- [4] Milan Sonka, Vaclav Hlavac, Roger Boyle, "Image Processing, Analysis and Machine Vision". *Chapman & Hall Computing*
- [5] Demas Sanger, Takuya Asada, Hiedeki Haneishi and Yoichi Muiyake, "Facial Pattern Detection and Its Preferred Color Reproduction", *The 2nd IS&T/Sid Color Imaging conference*, pp. 149-152, 1994
- [6] Rafael C. Gonzalez, Richard E. Woods "Digital Image Processing", Addison-Wesley Publishing Company
- [7] 전준현, "초고속 멀티미디어 통신 서비스를 위한 화상 데이터 압축 기술 분석", 통신학회지, 제 13

- 권, 제 5호, pp.77-101, 1996년 5월
- [8] 정제창 역, "최신MPEG", 교보문고, 1997년
- [9] H.Gharavi and M.Mills, "Block Matching Motion Estimation Algorithms - New results", *IEEE Transactions on Circuits and Systems*, Vol. 37, No.5, May 1990, pp. 649-651.
- [10] M.J. Chen, L.G. Chen and T.D. Chieuh, "One-Dimensional Full Search Motion Estimation Algorithm for Video Coding", *IEEE Trans. On Circuits and Systems for Video Technology*. pp.405-509.
- [11] R. Srinivasan and K.R. Rao, "Predictive Coding Based on Efficient Motion Estimation", *IEEE Trans. Commun.*, vol. COM-33, no.7, Aug. 1982.
- [12] Feng, K. T. Lo, H.mehrpour and A.E. Darbowiak, "Adaptive Block Matching Motion Estimation Algorithm for Video Coding", *Electronics letters*, Vol 32, No 18, pp 1542-1543.
- [13] ISO/IEC JTC1/SC29/WG11/M3100, "MPEG-4 Video Verification Model version.9.1"

저 자 소 개



李 奉 鎬(正會員)

1997년 2월 한국항공대학교 전자공학과 졸업(학사). 1999년 2월 한국항공대학교 전자공학과 석사과정 졸업(공학석사). 1999년 7월~현재 한국전자통신연구원(ETRI) 연구원. 주관심분야는 영상신호처리, 디지털 방송, MPEG-4, 멀티미디어 시스템



郭 魯 尹(正會員)

1994년 2월 한국항공대학교 항공전자공학과 졸업(공학사). 1996년 2월 한국항공대학교 대학원 항공전자공학과(공학석사). 2000년 2월 한국항공대학교 대학원 항공전자공학과(공학박사) 1998년 3월~2000년 2월 유한대학 정보통신과 겸임교수. 2000년 3월~현재 : 천안대학교 정보통신학부 전임강사 주관심 분야는 정지 영상 및 동영상 압축, 칼라 영상 처리 및 검색, 컴퓨터 비전, 영상 디스플레이

崔 敬 洙(正會員)

1980 경북대학교 전자공학과 학사. 1984 연세대학교 전자공학과 공학석사. 1980~현재 한국전자통신연구원(ETRI) 책임연구원, 주관심분야는 원격 교육 시스템, 위성 멀티미디어 시스템



黃 秉 元(正會員)

1972년 2월 한국항공대학교 항공전자공학과. 1993년 8월~1994년 8월 동경대학 전기공학과 교환교수. 1985년 8월~현재 한국항공대학교 항공전자공학과 교수. 주관심 분야는 영상처리 및 영상인식, 동영상 신호처리, 교통류 제어