

시간축 서브밴드 해석을 이용한 적응적 움직임 추정에 관한 연구

正會員 임 중 곤*, 정 재 호**

An Adaptive Motion Estimation Based on the Temporal Subband Analysis

Zung Kon Yim*, Jae Ho Chung** *Regular Members*

본 연구는 95년도 인하대학교 연구비 지원에 의하여 수행되었음.

요 약

움직임 벡터 추정은 MPEG과 같은 고화질용 동영상 압축기의 핵심 기술이다. 본 논문에서는 MPEG과 같은 비디오 부호화기에서 적용할 수 있는 새로운 움직임 추정 기법을 제안한다. 제안된 시간적으로 적응적인 움직임 추정기는 다음의 다섯가지 기능으로 요약된다. 시간축 서브밴드 해석, 시간축 정보추출, 장면전환 추적, 장면대체 그리고 시간적으로 적응적인 움직임 추정기법이다. 위의 모든 기능적 요소들은 시간축 서브밴드 해석에 기초를 둔다.

본 연구에서는 서브밴드 분해 및 합성 기법중 해석부분을 동영상 시퀀스의 시간축에 적용하였고, 시간적 활동량 분포 및 평균 시간적 활동량을 정의하였고 또한 앞에서 정의된 결과를 이용하는 시간적응적 블록정합 알고리즘, 장면전환 추적 알고리즘, 장면대체 알고리즘을 제안하였다.

시간적응적 움직임 추정 알고리즘은 블록 정합 영역을 제어할 수 있다. 운동체의 시간적 활동량 분포에 따라서 움직임 추정시 탐색영역을 비확률적으로 할당할 수 있다. 제안된 장면전환 추적 알고리즘과 장면대체 알고리즘은 급격히 전환된 장면에 대한 불필요한 움직임 예측에 사용되는 계산을 방지할 수 있다.

모의실험 결과, 제안된 방법이 기존의 MPEG-전역탐색 방법의 40%정도의 블록정합 시도로 보다 좋은 화질을 보였다. 또한 제안된 장면전환 추적기법을 사용하여 추가적인 계산없이도 실험에 사용된 임의로 조합된 영상 내의 장면전환을 모두 추적하였다.

*KBS 기술연구소기 연구실

**인하대학교 전자공학과

論文番號:95322-0916

接受日字:1995年 9月 16日

ABSTRACT

Motion estimation is one of the key components for high quality video coding. In this paper, a new motion estimation scheme for MPEG-like video coder is suggested. The proposed temporally adaptive motion estimation scheme consists of five functional blocks: Temporal subband analysis (TSBA), extraction of temporal information, scene change detection (SCD), picture type replacement (PTR), and temporally adapted block matching algorithm (TABMA). Here all the functional components are based on the temporal subband analysis.

In this paper, we applied the analysis part of subband decomposition to the temporal axis of moving picture sequence, newly defined the temporal activity distribution (TAD) and average TAD, and proposed the temporally adapted block matching algorithm, the scene change detection algorithm and picture type replacement algorithm which employed the results of the temporal subband analysis.

A new block matching algorithm TABMA is capable of controlling the block matching area. According to the temporal activity distribution of objects, it allocates the search areas nonuniformly. The proposed SCD and PTR can prevent unavailable motion predictions for abrupt scene changes.

Computer simulation results show that the proposed motion estimation scheme improves the quality of reconstructed sequence and reduces the number of block matching trials to 40% of the numbers of trials in conventional methods. The TSBA based scene change detection algorithm can detect the abruptly changed scenes in the intentionally combined sequence of this experiment without additional computations

I. 서 론

80년대에 들어서 널리 알려진 뉴미디어라는 용어는 디지털 기술의 급속한 발전과 함께 전개 양상이 다양화, 가속화되면서 사회전반에 큰 영향을 미치게 되었다. 뉴미디어는 영상, 음성, 음향 등의 디지털화된 막대한 데이터를 고속으로 송수신해야 한다는 중대한 역할을 감당해야한다. 그러나 일반적으로 영상 신호와 음성신호를 아날로그에서 디지털 신호로 변화하면 정보량이 크게 증가한다. 이러한 정보를 송수신할 때 전송 채널의 유효이용이라는 점에서는 압축률을 높여야하지만, 압축률을 높이면 높일수록 품질이 저하되기 마련이다. 또한 실제 송수신 시에는 전송 채널상의 잡음, 전파의 경우는 반사와 비직선성 등에 의해 수신측에서 부호오차가 발생하므로 오류정정을 위한 정보를 원래의 데이터에 추가로 첨가시켜야 올바른 데이터의 전송이 이루어질 수 있다.

디지털 비디오는 영상회의, 영상전화등 통신 분야 뿐만 아니라 컴퓨터 및 가전업계와 같은 넓은 응용분야에서 이용된다. 정보처리 업계들의 서로다른 응용분야들간의 기술공유 관점에서 ISO(International Orga-

nization for Standardization)는 디지털 저장매체용의 비디오 및 관련 오디오 표준개발을 시작했고 그러한 표준화 노력이 MPEG (Moving Picture Expert Group)을 탄생시켰으며, ISO-IEC/JTC1/SC29/WG11에 공식 위원회를 갖게되었다.

MPEG비디오 부호화 방식은(이하 MPEG)는 임의 접근(Random Access)과 고압축률을 위하여 I-화면, P-화면, B-화면의 구조로 동영상상을 부호화한다. 그 중에서 P-화면과 B-화면은 기준 프레임으로부터 움직임 추정 및 보상기법을 사용하여 현재 프레임을 예측하여 부호화하는 영상이다. 여기에 사용되는 움직임 추정 방법은 블록정합방식(BMA: block matching algorithm)을 주로 사용하게 되는데 단방향, 또는 양방향 움직임 추정을 위하여 매크로 블록단위의 정합을 정해진 탐색영역에 대하여 시도한다. 이러한 움직임 벡터 추정 방법은 상당한 계산량에도 불구하고 움직임벡터를 오추정하는 경우가 자주 있다. 또한, 연속된 영상 중에 급격한 장면 전환이 발생하는 경우에 대한 움직임 추정 및 보상 기법은 거의 무의미하고 할 수 있다.

위와 같은 문제점을 해결하기 위하여, 본 논문에서

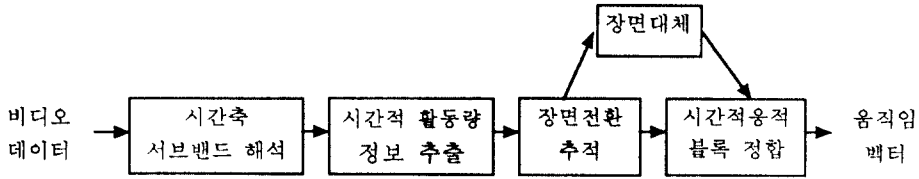


그림 1. 적응적 움직임 추정기의 블록도.

는 동영상 입력 신호로부터 실제의 움직임에 가까운 변위벡터를 추출할 수 있는 적응적인 움직임 추정기법을 제안한다. 본 논문에서 제안된 영상신호 부호화 시스템은 그림 1과 같이 시간축 서브밴드 해석(TSBA: Temporal Subband Analysis)[1][2], 시간적 활동량 정보 추출(Extraction of Temporal Activity Information), 장면전환추적 및 대체(SCD & PTR: Scene Change Detection & Picture Type Replacement), 시간적응적 블록정합(TABMA: Temporally Adapted Block Matching Algorithm)등, 다섯가지 기능으로 구성된다.

최근 동영상 부호화에 대한 연구에서 서브밴드 부호화(SBC: subband coding) 방법들이 제안되고 있다. 서브밴드 부호화에서는 1장의 영상을 여러개의 주파수 대역들로 분할하여 각 대역 특성에 따라 부호화하고 다시 합성한다. 이러한 서브밴드 부호화는 영상 전체에 대하여 변환하여 부호화하므로 저전송률의 블록 변환 부호화에서 나타나는 블록효과가 발생하지 않는 장점이 있다. 한 장의 영상에 대하여 서브밴드 부호화하는 방법을 공간축 서브밴드 부호화 (spatial subband coding)라 하는데 반면, 동영상내 여러장의 영상에 대하여 시간축에 따라서 서브밴드 부호화하는 방법을 시간축 서브밴드 부호화(temporal subband coding)라 한다. 특별히 두가지 방법을 다 사용하는 동영상 서브밴드 부호화 방법을 3차원 서브밴드 부호화 (3-dimensional subband coding)라 한다. 본 논문에서는 시간축 서브밴드 부호화 방법중 서브밴드 해석 기법만 이용하여 연속된 시퀀스 내에 존재하는 시간적 활동량을 정의하여 구하고 그 분포를 기초로하여 시간적응적 블록정합 알고리즘과 장면전환 및 장면대체 알고리즘을 제안하였다.

본 논문에서는 우선 실제 운동체의 변위에 가까운 움직임벡터를 효과적으로 찾아내기 위하여 시간적응적 블록정합 알고리즘(TABMA)를 제안한다. 제안된

TABMA는 기존의 MPEG에서 획일적으로 적용하는 탐색영역의 크기와는 달리, 탐색영역을 각 매크로블록의 시간적 활동량에 비례하게 할당하여 적은 계산량으로 정확한 변위벡터를 찾아낸다. 기존의 MPEG에서 사용하는 전역탐색방법은 획일적으로 할당된 영역에 대하여 블록정합을 시도하여 단순히 흑백성분에 대한 평균절대오차만 최소화하는 이동변위를 움직임벡터로 결정하기 때문에, 실제 변위와는 다른 움직임벡터를 결정할 수 있다. 또한 Miss America 영상의 배경에서와 같이 시간적인 변화량이 거의 없는 영역에 대해서도 같은 방법으로 전역탐색을 시도하므로 계산량의 관점에서 비효율적이다.

본 논문에서는 한 동영상 시퀀스내에 급격한 장면전환이 발생할 경우 간단한 계산만으로 전환된 장면을 추적하고, 또한 불필요한 장면 예측에 대한 계산량을 감축시키기 위하여 장면전환 추적기와 장면대체 알고리즘을 제안한다. 제안된 알고리즘들은 모두 Johnston이 설계한 FIR 필터뱅크를 이용하여 시간축을 따라서 분할된 서브밴드에서 추출한 시간적 활동량정보로부터 구현된다[3]. 모의실험 결과, 제안된 방식은 블록정합을 시도한 횟수와 부호화된 영상의 정량적 화질에서 기존의 MPEG 전역탐색 기법에 비하여 모두 우수한 성능을 보였다.

논문의 제 2절에서는 제안된 알고리즘의 기본이 되는 시간축 서브밴드 해석을 설명하였고, 제 3절과 제 4절에서는 제안된 방식들을 설명하였다. 또한 제 5절에서는 이 제안된 방식을 여러가지 시험영상에 적용하여 얻은 실험결과를 제시하였다. 마지막으로 본 논문의 결론이 제 6절에 기술되었다.

II. 시간축 서브밴드 해석

일반적으로 시간축 서브밴드 해석 및 합성 기법은

공간 서브밴드 해석 및 합성 기법과 함께 동영상 부호화에 사용되어 3차원 서브밴드 해석 및 합성 기법으로 많이 연구되고 있다. 하지만 본 논문에서는 서브밴드 해석 기법을 영상시퀀스의 시간축에 적용하고 본 논문에서 정의한 시간적 활동량 분포 및 평균 활동량을 추출하여 제안된 움직임 추정을 하는데 이용한다.

MPEG 비디오 부호화에서는 일반적으로 12~16장의 연속된 프레임을 GOP(Group of Pictures) 계층으로 묶어 임의접근(Random Access)의 단위가 되도록 하였다. 본 논문에서는 GOP내의 영상 중에서 실질적으로 움직임 추정 및 보상의 단위가 되는 4 프레임을 sub-GOP라 정의하고, 새로 정의된 sub-GOP내의 4 프레임은 시간축 서브밴드 해석의 단위로 사용된다.

본 논문에서 사용한 시간축 서브밴드 해석 시스템은 시간적으로 연속하는 4 프레임을 4개의 주파수 대역으로 분할한다. 시간축 서브밴드 해석 과정은 다음과 같고, 수식은 식(1)~(4)에, 시간축 서브밴드 해석 시스템의 전체 블록도는 그림 2에 나타내었다[4][5].

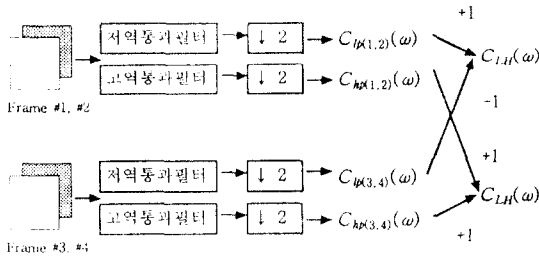


그림 2. 시간축 서브밴드 해석.

시간축 서브밴드 해석 과정

- 과정 1: sub-GOP내의 4개의 프레임중 앞의 두 프레임에 대하여 저역 및 고역 통과 필터링한다.
- 과정 2: 생성된 두 개의 대역에 대하여 부표본화를 수행하여 $C_{lp(1,2)}(w)$ 와 $C_{hp(1,2)}(w)$ 대역으로 분할한다.
- 과정 3: sub-GOP내의 나머지 두 프레임에 대하여 과정 1과 과정 2를 반복하여 $C_{lp(3,4)}(w)$ 와 $C_{hp(3,4)}(w)$ 대역으로 분할한다.
- 과정 4: $C_{lp(3,4)}(w)$ 와 $C_{lp(1,2)}(w)$ 의 차로부터 $C_{LH}(w)$ 대역을 구하고 $C_{hp(1,2)}(w)$ 와 $C_{hp(3,4)}(w)$ 의

합으로부터 $C_{LL}(w)$ 대역을 구한다.

$$C_{lp(1,2)}(w) = \frac{1}{2} [H_{lp}(\frac{w}{2}) X_{(1,2)}(\frac{w}{2}) + H_{lp}(\frac{w}{2} + \pi) X_{(1,2)}(\frac{w}{2} + \pi)] \quad (1)$$

$$C_{hp(1,2)}(w) = \frac{1}{2} [H_{hp}(\frac{w}{2}) X_{(1,2)}(\frac{w}{2}) + H_{hp}(\frac{w}{2} + \pi) X_{(1,2)}(\frac{w}{2} + \pi)] \quad (2)$$

$$C_{lp(3,4)}(w) = \frac{1}{2} [H_{lp}(\frac{w}{2}) X_{(3,4)}(\frac{w}{2}) + H_{lp}(\frac{w}{2} + \pi) X_{(3,4)}(\frac{w}{2} + \pi)] \quad (3)$$

$$C_{hp(3,4)}(w) = \frac{1}{2} [H_{hp}(\frac{w}{2}) X_{(3,4)}(\frac{w}{2}) + H_{hp}(\frac{w}{2} + \pi) X_{(3,4)}(\frac{w}{2} + \pi)] \quad (4)$$

과정 4로부터 얻어지는 두 대역은 sub-GOP 내의 4 프레임에 대한 시간축 정보를 포함하고 있다. 그 내용을 식 (5)와 (6)에 나타내었다.

$$C_{LH}(w) = \frac{1}{2} [C_{lp(1,2)}(w) - \overline{C_{lp(3,4)}(w)}] \quad (5)$$

$$C_{LL}(w) = \frac{1}{2} [C_{hp(1,2)}(w) + C_{hp(3,4)}(w)] \quad (6)$$



그림 3. 시간적으로 서브밴드 필터링된 Flower Garden의 $C_{LH}(w)$ 대역 영상.

III. 제안된 움직임 추정 알고리즘

본 논문에서는 각 매크로블록 좌표 범위의 시간적 활동량을 구하여 실질적으로 필요한 구간 만큼의 탐

색영역을 할당하고, 급격한 장면 전환의 발생시 간단한 계산만으로 전환된 장면을 추적하여 불필요한 프레임 예측을 배제시킴으로써 프레임 예측 이득의 향상과 계산량 감축에 기여할 수 있는 알고리즘을 제안한다.

3.1 시간적 활동량 분포

본 논문에서는 시간적 활동량과 평균 시간적 활동량을 각각 TAD (Temporal Activity Distribution)와 TAD_{AVG} 로 정의하였다. 여기서 TAD 는 sub-GOP 내의 4 프레임의 시간적 활동량 분포를 나타내고, TAD_{AVG} 는 sub-GOP 전체의 시간적 활동량을 나타낸다. 식 (7)과 (8)에서 TAD 와 TAD_{AVG} 에 대한 식을 나타내었다.

$$TAD_{mb}(k, l, band) = \frac{N(k+1) \times 16 - 7M(l+1) \times 16 - 7}{\sum_{i=k \times 16 + 8}^{N(k+1) \times 16 - 7} \sum_{j=l \times 16 + 8}^{M(l+1) \times 16 - 7} |C_{band}(i, j)|} \quad (7)$$

$$TAD_{AVG, sub-GOP}(band) = \frac{1}{M \times N} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} TAD_{mb}(k, l, band) \quad (8)$$

여기서 M 은 한 주사선 위의 화소의 수이고, N 은 한 프레임 내의 주사선의 수이고, $band$ 는 시간축에 대하여 분해된 서브밴드 인덱스 $\{hp(1, 2), hp(3, 4), LH, HL\}$ 이다. 또한 $TAD_{mb}(k, l, band)$ 는 $band$ 라는 인덱스로 표시되는 서브밴드의 수직으로 k 번째, 수평으로 l 번째에 위치하는 매크로블록의 시간적 활동량을 나타낸다. TAD 가 해당 서브밴드 내의 모든 매크로블록에 대하여 계산이 되면 각 서브밴드의 평균 시간적

활동량 $TAD_{AVG, sub-GOP}(band)$ 은 식 (8)에 의하여 구한다. 여기서 $TAD_{AVG, sub-GOP}(band)$ 는 서브밴드 분해된 해당 sub-GOP의 평균 시간적 활동량을 나타낸다. 그림 4에 'Flower Garden' 영상에 대한 시간적 활동량 분포를 예로 나타내었다. 그림 4에 나타난 바와 같이 시간적인 관점에서 비활동적인 영역 예를들면 Flower Garden 영상에서 하늘 부분의 경우 시간적 활동량 레벨이 상당히 낮으며, 움직이는 나무 부분의 경우 상대적으로 높은 시간적 활동량 레벨을 보이고 있다.

3.2 시간적응적 블록정합 알고리즘

제안된 시간적응적 블록정합 알고리즘은 탐색영역의 크기를 시간적 활동량에 비례하도록 할당하여, 실제의 움직임에 가까운 움직임 벡터를 추정할 수 있도록 하는데 목적이 있다. 예를들어, Miss America 영상의 배경 부분, 또는 Flower Garden 영상의 하늘 부분과 같이 시간적인 활동량이 적은 부분에 대해서는 아주 적은 탐색영역을 할당하고, 그 외의 시간적으로 활동량이 많은 부분에 대해서는 활동량에 비례하는 넓은 탐색영역을 할당하도록 한다.

본 논문에서는 시간적 활동량의 크기에 비례하도록 탐색영역을 적용시키기 위하여 식 (9)를 정의하였다. 여기서 $SA(k, l)$ 는 좌표 (k, l) 에 위치한 매크로블록이 움직임 벡터를 찾기 위하여 탐색하는 정방형 탐색영역의 한 변의 길이이다.

$$SA(k, l) = \begin{cases} SA_{MIN} + c \times TAD_{mb}(k, l, LH), & 0 \leq TAD_{mb}(k, l, LH) \leq threshold \\ SA_{MPEG}, & threshold \leq TAD_{mb}(k, l, LH) \end{cases} \quad (9)$$

여기서 SA_{MPEG} 은 MPEG 비디오 부호화에서 사용하는 일반적인 정방형 탐색영역의 한 변의 길이이다. $SA(k, l)$ 는 항상 SA_{MPEG} 보다 작거나 같은 값을 갖도록 하여 기존의 탐색영역 보다 커지지 않도록 하였고, c 는 $SA(k, l)$ 가 $TAD_{mb}(k, l, LH)$ 에 비례하여 할당됨을 나타내는 비례상수이다. 제안된 탐색영역에서는 후보 움직임 벡터의 수가 기존의 MPEG의 경우 보다 작아지므로 블록정합 시도횟수가 절감될 수 있다. 그리고 탐색영역이 기존의 방법보다 크지 않도록 제한한 이유는 기존의 MPEG 움직임 추정의 범위로

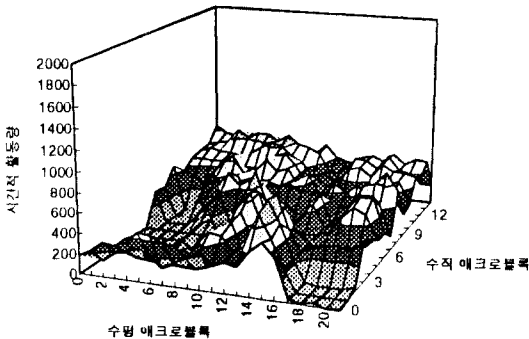


그림 4. Flower Garden 영상의 시간적 활동량 분포.

도 충분한 탐색이 가능하기 때문이다.

IV. 장면전환 추적과 장면대체 알고리즘

본 논문에서는 장면전환 추적과 장면대체 알고리즘을 제안하여, 움직임 추정시 불필요한 계산량을 최소한으로 줄이고, 실제의 움직임에 가까운 움직임벡터를 찾을 수 있도록 하였다. 제안된 장면전환 추적 및 장면 대체 알고리즘은 앞의 서브밴드 해석을 이용한다(2장 참조). 본 장에서 제안한 두 알고리즘은 앞서 계산된 시간적 활동량 분포와 평균 시간적 활동량을 사용하므로 추가로 소요되는 계산없이 구현할 수 있다.

4.1 장면전환 추적 알고리즘

MPEG 비디오 부호화 알고리즘은 예측부호화 방법을 사용하는데, 이러한 알고리즘에 있어서 시퀀스 내에 급격한 장면의 존재는 예측부호화의 효율을 급격히 저하시키는 원인이된다. 그러므로 부호화를 시작하기 전에 장면전환의 존재와 그 위치에 대한 정보를 가지고 있는 것은 중요하다.

영상의 시퀀스 중에 존재할 수 있는 장면전환의 문제는 크게 두가지로 분류된다. 첫째는, '해당 sub-GOP 내에 급격한 장면전환(ASC: Abrupt Scene Change)이 존재하는가?', 그리고 둘째는, 만약에 존재한다면, '몇 번째 프레임에서 부터 전환된 장면이 시작되는가?'이다.

일반적으로 sub-GOP내에 급격하게 전환된 장면이 있을 경우 sub-GOP내 4장의 프레임에 대하여 같은 좌표에 해당하는 화소가 갖는 값이 급격한 변화를 한다고 할 수 있다. 동일한 좌표의 연속된 화소값 4개에 대하여 고역통과 필터링을 하면 고주파 성분이 지배적이된다. 그러므로 장면전환을 포함하고 있는 sub-GOP의 $C_{LH}(\omega)$ 대역에 대하여 고주파 성분이 지배적이다. 여기서 $C_{LH}(\omega)$ 대역은 제 2 장의 그림 2에서 설명했듯이, sub-GOP의 처음 두 프레임을 저역통과 필터링한 $C_{lp(1, 2)}(\omega)$ 대역과 나중 두 프레임을 저역통과 필터링한 결과 대역 $C_{lp(3, 4)}(\omega)$ 대한 차 대역이다.

그림 5의 $C_{LH}(\omega)$ 는 sub-GOP내에 전환된 장면이 존재하는 '조합된 영상'의 LH 대역 서브밴드이다. 이 시퀀스는 두장의 'Flower Garden' 영상과 두장의 'Foot Ball' 영상으로 이루어져있다. 그러므로 sub-GOP의 3

번째 영상이 전환된 장면이다. 그림 6은 위의 '조합된 영상'에 대한 시간적 활동량 분포 $TAD_{mb}(k, l, LH)$ 이다. sub-GOP 내에 장면전환이 존재하는 경우와 장면전환이 존재하지 않는 경우를 비교해 보면, 전환된 장면이 없는 Flower Garden에 대한 그림 4의 $TAD_{mb}(k, l, LH)$ 는 화면상에서 움직임이 있는 부분에 대하여 주위의 정적인 영역에 비해 상대적으로 높은 값으로 분포되어 있다. 하지만 전환된 장면을 가진 '조합된 영상'에 대한 그림 6의 $TAD_{mb}(k, l, LH)$ 는 화면 전체적으로 장면전환이 없는 Flower Garden 영상에 비해 월등히 높은 시간적 활동량을 나타내고 있음을 알 수 있다.



그림 5. 시간적으로 서브밴드 필터링된 '조합된 영상'의 $C_{LH}(\omega)$ 대역 영상.

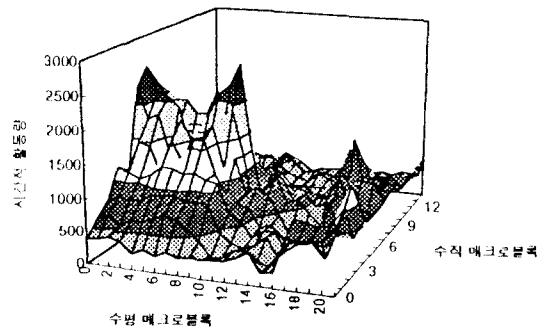


그림 6. '조합된 영상'의 시간적 활동량 분포.

위와 같은 현상으로부터 장면전환 추적을 위한 다음과 같은 과정을 생각할 수 있다. 먼저 $TAD_{AVG, sub-GOP}$

(LH)와 $TAD_{AVG, sub-GOP}(HL)$ 을 계산한다. 이 두 값의 산술평균이 사전에 정의된 문턱값 THR_{PIC} 보다 크면, 해당 sub-GOP내에 새로운 장면이 시작된다고 생각할 수 있다. 그리고나서 전환된 장면의 존재 여부를 확인하기 위하여 정해진 문턱값 THR_{LOW} 보다 작은 $TAD_{mb}(k, l, LH)$ 의 갯수를 검사하고, 그 갯수가 정해진 문턱값 THR_{NUM} 보다 작을 경우 해당 sub-GOP 내에 장면전환이 있다고 판정한다.

그 다음으로 장면전환이 있는 sub-GOP에 대하여 전환된 장면의 시작 위치를 알아내는 과정이 필요하다. 필터링된 $C_{hp(1, 2)}(\omega)$ 와 $C_{hp(3, 4)}(\omega)$ 두 대역은 각각 처음 두 프레임과 나머지 두 프레임에 대한 시간적 변화량에 대한 정보를 가지고 있다. $TAD_{AVG, sub-GOP}(hp(1, 2))$ 와 $TAD_{AVG, sub-GOP}(hp(3, 4))$ 의 산술평균이 정해진 문턱값 $THR_{POSITION}$ 보다 작은 경우엔 세번째 프레임에서 전환된 장면이 시작한다고 결정한다. 그렇지 않고 $TAD_{AVG, sub-GOP}(hp(1, 2))$ 가 $THR_{POSITION}$ 보다 큰 경우에는 두 번째 프레임에서 전환된 장면이 시작한다고 결정하고, $TAD_{AVG, sub-GOP}(hp(3, 4))$ 가 $THR_{POSITION}$ 보다 큰 경우에는 네 번째 즉, sub-GOP의 마지막 프레임에서 전환된 장면이 시작한다고 결정한다. 다음에 그 일관된 알고리즘을 소개하였다.

장면전환 추적 알고리즘(SCD)

```

if {  $\frac{1}{2} (TAD_{AVG, sub-GOP}(LH) + TAD_{AVG, sub-GOP}(HL)) \geq THR_{PIC}$  }
and if { Number of [  $TAD_{mb}(k, l, LH) < THR_{LOW}$  ]  $\leq THR_{NUM}$  }
    the SCENE IS CHANGED.
if { SCENE IS CHANGED }
if {  $\frac{1}{2} (TAD_{AVG, sub-GOP}(hp(1, 2)) + TAD_{AVG, sub-GOP}(hp(3, 4))) \geq THR_{POSITION}$  }
    then SCENE IS CHANGED AT THIRD FRAME.
else if {  $TAD_{AVG, sub-GOP}(hp(1, 2)) > THR_{POSITION}$  }
    then SCENE IS CHANGED AT FOURTH FRAME.
else SCENE IS NOT CHANGED.
    
```

4.2 장면대체 알고리즘

기존의 MPEG은 장면전환의 발생 여부와는 상관 없이 정해진 방향에 대한 프레임 예측을 시도하기 때문에 영상 시퀀스 내에 전환된 장면이 존재하면 블록

정합을 위한 계산량의 소비에도 불구하고 예측이득이 거의 없게된다. 본 논문에서는, 영상 시퀀스 내에 장면전환이 발생할 경우, 프레임 예측에 필요한 방향에 대해서만 블록정합을 수행하여 불필요한 블록정합을 사전에 방지하도록 sub-GOP의 마지막 기준 프레임인 P-화면을 움직임 추정을 하지 않는 유사 I-화면으로 교체하는 장면대체 알고리즘을 제안한다. 일반적으로 MPEG은 I, B, B, P 또는 P, B, B, P구조의 움직임 추정 방식을 사용하는데, 그림 7에서 보듯이 sub-GOP내에 전환된 장면이 추적되면, 제안된 방식에서는 sub-GOP의 마지막 화면을 블록정합하지 않는 INTRA 타입과 유사한 화면으로 대체시킨 후, 프레임 예측을 필요한 방향에 대해서만 수행하도록 한다.

본 논문에서는 P-화면으로 부호화해야하는 영상을 I-화면과 유사한 화면으로 처리함으로써 발생할 수 있는 버퍼의 오버플로우를 방지하기 위하여 원래의 P-화면으로 부호화할 때보다 5% 많은 양의 비트수를 할당하여 I-화면로 부호화하게 하였다. 그 이유는 sub-GOP의 마지막 프레임은 항상 프레임 예측에 대한 기준으로 사용되므로 마지막의 기준 프레임의 화질을 좋게하여 양방향 예측시 예측 이득을 높여 추가로 할당된 5%의 비트수를 보상할 수 있도록 하기 위함이다. 급격한 장면전환이 존재하는 sub-GOP에 대하여 제안된 장면대체 알고리즘을 사용할 경우, 화질

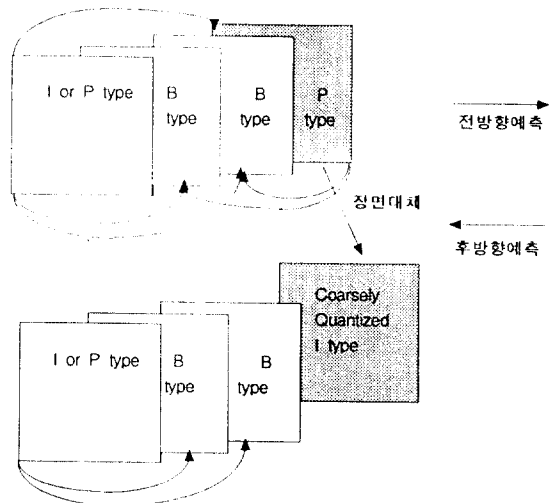


그림 7. 장면전환이 발생했을 경우의 장면대체 알고리즘.

은 보다 양호하게 유지하면서 기존의 방법보다 60% 정도의 블록정합 시도횟수를 절감할 수 있다.

V. 모의실험 및 결과

본 연구에서는 제안된 알고리즘에 대하여 세 개의 서로 다른 시험영상을 사용하였다. 시험영상은 다음과 같다. *Flower Garden* 영상 16 프레임(352×240), *Miss America* 영상 30 프레임 (352×288), 그리고 의도적으로 편집된 '조합된 영상' 16 프레임 (352×240). 여기서 '조합된 영상'은 5개의 전환된 장면을 포함하도록 의도적으로 편집을 하였고, 특별히 장면전환 추적과 장면대체 알고리즘의 성능을 평가하기 위하여 사용된 시험영상이다. 본 논문에서는 제안된 알고리즘을 기존의 MPEG에서 사용하는 전역탐색법과 성능을 비교하였다. 본 장에서 언급되는 비트율은 부호화기에서 출력되는 실제의 출력 비트율이다. 그림 8-10에서는 0.5Mbps에서 1.5Mbps까지 다양한 비트율에 대한 복원된 신호대 잡음비를 비교하고, 특히 *Miss America* 영상에 대해서는 0.1Mbps에서 0.45Mbps의 비트율에 대하여 동일한 성능을 비교하였다.

성능평가 곡선은 제안된 알고리즘이 기존의 전역탐색 방법에 비하여 좋은 성능을 나타내고 있음을 보이고 있다. 특별히 매우 작은 비트율과 또한 시간적으로 활동량이 적은 *Miss America*와 같은 영상에 대하여 상당히 좋은 성능을 보이고 있음을 알 수 있다.

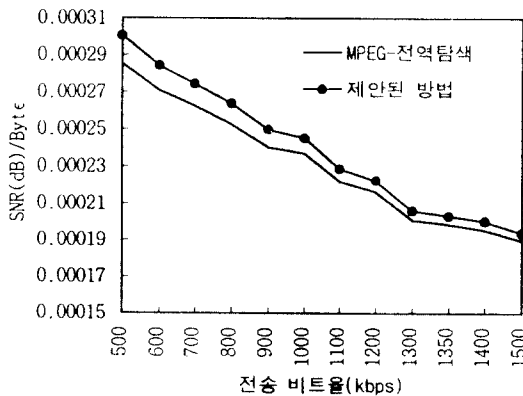


그림 8. 'Flower Garden' 영상에 대한 화질 곡선

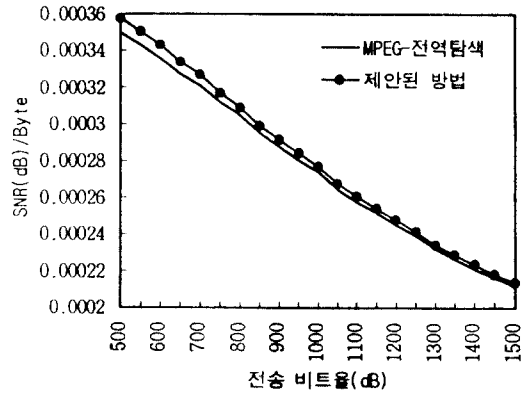


그림 9. '조합된 영상'에 대한 화질 곡선.

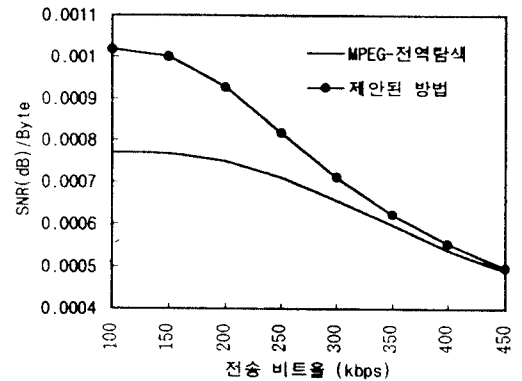


그림 10. 'Miss America' 영상에 대한 화질 곡선.

표 1로부터 제안된 알고리즘과 기존의 전역탐색법이 블록정합 횟수의 관점에서 비교되고 있으며 제안된 알고리즘으로 평균 60%이상의 블록정합 횟수를 절감하고 있음을 알 수 있다. 특별히 *Miss America* 영상에 대해서는 약 12%정도의 블록정합만으로 우수한 화질을 나타낼 수 있었는데 그 이유는 시험영상 자체가 화면내 움직임이 별로 없는 영상이기 때문이다. 또한 조합된 영상의 경우 제안된 장면전환 추적

표 1. sub-GOP당 평균 블록정합 횟수

	Flower Garden	조합된 영상	Miss America
제안된 알고리즘	124,514	138,066	48,863
MPEG 전역탐색법	330,750	330,750	402,625
비율	37.65%	41.74%	12.14%

알고리즘과 장면대체 알고리즘은 급격히 변환된 장면을 모두 찾아 내었으며 그 결과로 전체적인 계산량 감소에 기여하였다.

VI. 결 론

본 논문에서는 시간축 서브밴드 해석 방법을 이용하여 새로운 움직임 추정 기술을 개발하였다. 제안된 알고리즘은 시간축 서브밴드 해석으로부터 유도된 움직이는 물체에 대한 시간적 활동량을 이용한다. 기존의 MPEG 등의 동영상 압축알고리즘에서는 움직임 추정 및 보상 방법을 사용하는데, 이러한 움직임 추정을 위한 탐색영역의 크기가 본 논문에서 제안된 시간적으로 적응적인 블록정합 알고리즘에 의하여 산출이된다. 또한 본 논문에서 제안된 장면전환 추적 알고리즘과 장면대체 알고리즘을 이용하여 동영상 내에 존재하는 전환된 장면의 시작 프레임을 추적하고 전환된 장면에 대하여 효과적인 압축을 할 수 있도록 하였다.

제안된 방법은 기존의 MPEG 등의 동영상 부호화 기술에 비하여 적은 계산량으로 좋은 화질의 영상으로 부호화 할 수 있다. 특별히 낮은 비트율에서 그리고 Miss America와 같은 시간적으로 활동량이 적은 영상에 대하여 매우 좋은 성능을 보인다. 계산량의 관점에서 볼 때 제안된 알고리즘은 일반적인 비디오 영상에 대하여 전형적인 움직임 추정기에 비해 약 40% 정도의 블록정합 횟수만으로 양질의 화질을 얻을 수 있게 하였다. 특별히 시간적으로 활동량이 적은 영상에 대하여는 약 12% 정도의 블록정합으로도 훨씬 좋은 부호화 화질을 얻을 수 있었다.

참 고 문 헌

1. J. R. Ohm, "Three-Dimensional Subband Coding with Motion Compensation," *IEEE Trans. Image Processing*, vol. 3, NO. 5, pp. 559-571, Sept. 1994.
2. J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pp. 291-294 1980.
3. O. Johnsen, O. V. Shentov and S. K. Mitra, "A

technique for the efficient coding of the upper bands in subband coding of images," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pp. 2097-2099, April, 1990.

4. R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1983.
5. H. Garavi and A. Tabatabai, "Sub-band coding of monochrome and color images," *IEEE Trans. Circuits and Systems*, vol. 35, pp. 207-214, Feb. 1988.



임 중 곤(Zung Kon Yim)정회원

1969년 6월 27일생

1994년 2월:인하대학교 공과대학 전자공학과(학사)

1996년 2월:인하대학교 공과대학 전자공학과(석사)

1996년 1월~현재:KBS 기술연구소 기기연구실 연구원

※주관심분야:비디오 신호처리, 오디오 신호처리, Digital Audio Broadcasting



정 재 호(Jae Ho Chung) 정회원

1982년:美國 University of Maryland (공학사)

1984년:美國 University of Maryland (공학석사)

1990년:美國 Georgia Institute of Technology(공학박사)

1984년~1985년:美國 국방성산하 해군연구소, 신호처리실, Electronic Engineer

1991년~1992년:美國 AT&T Bell 연구소, 음성 신호처리 연구실, 연구원(Member of Technical Staff)

1992년~현재:인하대학교, 전자공학과, (현)부교수

1995년~현재:한국전자통신연구소, 자연어처리연구실 초빙 연구원