

---

# 움직이는 물체의 고속 검출이 가능한 디지털 감시 시스템

김선우\* · 최연성\* · 박한엽\*

Digital Surveillance System with Fast Detection of Moving Object

Sun-Woo Kim\* · Yeon-sung Choi\* · Yeop-Han Park\*

---

이 논문은 2000년도 대학기초 연구비를 지원받았음

---

## 요 약

본 논문에서는 현재 사용하고 있는 아날로그 방식의 감시 시스템에서 발생하는 자원 낭비와 효율성 저하 문제를 극복하기 위해 MPEG-2 국제 표준을 이용한 영상 데이터의 인코딩 기법과 압축 기법을 적용한 새로운 방식의 디지털 감시 시스템의 설계 및 구현 기법에 대해 기술하였다. 또한, 기존의 디지털 감시 카메라 시스템에서 적용하지 못한 고속 움직임 추정 알고리즘들을 적용하여, 보다 실시간적인 시스템을 제안한다. 고속 움직임 추정 알고리즘들은 MPEG-2 비디오 인코딩을 위해서 제안되었다. 이 알고리즘들은 블록 매칭 기법과 그래디언트 기법의 혼합적인 사용에 기반한다. 또한 MPEG-2 비디오 데이터를 직접적으로 사용하여 움직이는 물체를 추출하는 방법을 제안하였다. 제안된 방법은 매우 간단하고 종래의 물체 검출 방법들 보다 계산 시간이 훨씬 덜 요구된다. 본 논문에서 제안한 시스템은 특정한 하드웨어 없이 단지 소프트웨어만으로 영상데이터를 실시간으로 인코딩 및 디코딩 한다.

## ABSTRACT

In this paper, since we currently using surveillance system of analog type bring about waste of resource and efficiency deterioration problems, we describe new solution that design and implementation to the digital surveillance system of new type applying compression techniques and encoding techniques of image data using MPEG-2 international standard. Also, we proposed fast motion estimation algorithm requires much less than the convectional digital surveillance camera system. In this paper a fast motion estimation algorithm is proposed the MPEG-2 video encoding. This algorithm is based on a hybrid use of the block matching technique and gradient technique. Also, we describe a method of moving object extraction directly using MPEG-2 video data. Since proposed method is very simple and requires much less computational power than the conventional object detection methods. In this paper we don't use specific H/W and this system is possible only software encoding, decoding and transmission real-time for image data.

## 키워드

MPEG-2, Surveillance system, Fast Motion estimation algorithm

---

\*군산대학교 전자정보공학부  
접수일자 2001.2.24

### 1. 서 론

최근 들어 컴퓨터를 이용한 데이터 처리기술 및 통신 기술이 급격히 발전함에 따라 정지영상, 동영상, 음성 등의 다양한 데이터를 이용한 멀티미디어 시스템이 개발 및 이용이 증가하고 있다.

그 중에서도 이미지나 동영상 그리고 음성 등의 압축 기술이 급속도로 발전함에 따라서 대량의 영상 데이터를 고화질의 디지털 영상으로 압축하여 저장할 수 있게 되었다. 이에 따라 카메라와 디지털 영상 인코딩 및 압축 기술은 영상 전화, 영상회의, 영상교육, 감시카메라 등의 첨단 분야에 널리 적용되어지고 있다.

움직임 추정은 움직임 보상 예측에 이용되는 움직임 벡터들을 생성하는데 필수적인 것이고 MPEG에 포함된 중요한 코딩 기법들 중의 하나이다. MPEG-2 비디오 표준[1]은 움직임 보상을 블록 영역에 제한하는 반면, 블록 벡터들이 추정되는 방법은 전적으로 인코더 설계에 맡긴다.

따라서 움직임 추정을 얼마나 빠르게 하느냐에 따라서 실시간이 가능한 감시 시스템을 구현할 수 있다.

본 논문에서는 고속으로 움직임을 추정하는 움직임 추정 알고리즘과 고속으로 물체를 검출할 수 있는 다음과 같은 알고리즘을 제안한다.

고속 움직임 추정 알고리즘은 블록 매칭 기법과 적응적 혼합 매칭 기법의 두 가지 기법의 혼합을 통해서 빠른 속도의 움직임 추정 알고리즘을 만들어 낼 수 있다. 이 혼합 기법은 MPEG-2 비디오 인코딩에서 거의 실시간에 가까운 소프트웨어를 구현하기 위해서 사용되어질 수 있다.

본 알고리즘에서, 부호화된 비트스트림에서 직접적으로 추출될 수 있는 움직임 정보같은 부호화된 정보만을 이용한다.

제안한 시스템에서는 특정한 하드웨어 없이 소프트웨어만으로 영상데이터를 실시간으로 인코딩 및 디코딩 할 수 있으며, 물체의 움직임을 추정하여 움직임 검출이 가능할 뿐만 아니라 인터넷을 이용한 원격지 감시 기능과 장거리 영상 전송이 가능한 IP 기반의 MPEG-2 디지털 감시 카메라 시스템의 개발에 대해 기술하였다.

본 논문의 구성은 다음과 같다. 우선 2장에서는 MPEG-2 인코딩과 움직임 추정과 실시간 전송을 위한 RTP/RTCP 프로토콜에 대해서 간략하게 소개를 하고, 3장에서는 본 논문에서 개발한 감시 시스템의 구성도에 대해서 기술한다. 4장에서는 본 논문에서 제안한 고속 움직임 검출 알고리즘에 대해서, 5장에서는 제안된 알고리즘에 대한 성능 분석 결과에 대해 기술한다. 마지막으로 6장에서는 결론 및 향후 연구 방향에 대해 기술한다.

### II. MPEG-2에서의 움직임 추정

#### 1. MPEG-2 인코더

MPEG 표준은 가장 보편적으로 이용되는 뛰어난 압축률을 가지는 동영상에 관한 국제적인 압축 표준이다. 기본적인 MPEG-2 인코더 알고리즘은 그림 1과 같다.

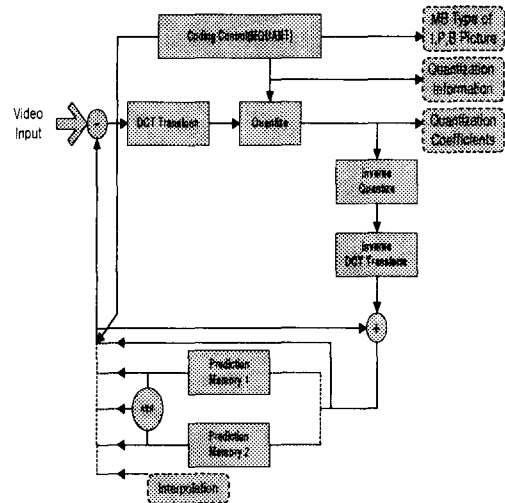


그림 1. MPEG-2 인코더 블록도  
Figure 1. MPEG-2 Encoder Block diagram

그림 1에서 비디오 신호 입력은 DCT, 양자화, 반화소 움직임 보상, 예측 부호화를 통해 압축되어지고 VLC를 통해서 인코딩된다. 이는 다시 반

대의 과정인 역양자화, IDCT, 예측 복호화를 거쳐 화면에 출력된다.

MPEG 표준에서 제안된 동영상 압축 기술의 영상 프레임 압축 방법은 크게 프레임내(intra frame) 압축과 프레임간(inter frame) 압축으로 나눌 수 있다[2][3].

프레임내 압축이란 시간적으로 서로 인접한 프레임 간의 연관 정보에 관계없이 자체적으로 압축을 하는 방법이다.

프레임간 압축은 연속된 프레임 사이에 존재하는 영상정보는 서로 연관되어 있다는 특징을 이용하여 인접한 프레임 사이의 영상 정보 중에서 변경된 정보만으로 프레임을 구성하여 압축하는 방법을 말한다. 따라서 프레임간 압축기법은 프레임 내의 압축 보다 효과적인 압축률을 제공할 수 있다.

순수 MPEG-2에서 제안된 부호화 알고리즘은 영상 데이터를 압축하기 위해 필요한 프레임 데이터간의 시간적 중복을 얻기 위해서 블록 기반의 움직임 보상을 사용하며, 공간적 중복을 얻기 위해서는 압축영역 변환 기법을 사용하고 있으며, 움직임 보상 기법에서는 예측 및 보간 기법이 이용되고 있다[4].

## 2. 움직임 추정 알고리즘

움직임 추정 기법에서는 각각의 화소가 화소값의 변화없이 위치만 움직인다고 가정하여 현재 프레임의 화소값을 추정하며, 추정된 화소값과 실제 화소값의 화소 차이와 움직임 정보만을 전송한다.

움직임 추정 및 보상 기법으로는 화소 순환 기법(Pixel-Recursive)과 블록 매칭 기법(Block matching), 그래디언트 기법(Gradient)으로 분류된다. 화소 순환 알고리즘은 연산량이 많고 복잡하며, 수렴하지 않는 문제가 발생하기 때문에 잘 사용되지 않는다. 블록 매칭 기법은 수행 시간이 비교적 적게 소요되고, 하드웨어적으로 구현이 용이하기 때문에 화소 순환 기법보다는 많이 사용되어 진다. 블록 매칭 기법은 현재 픽처와 참조 픽처 사이에 차분값을 최소화하는데 기반한

것으로, 블록 단위로 이전 프레임의 탐색 영역 내에서 매칭 오차가 최소인 위치를 찾아 그 위치를 해당 블록의 움직임 벡터로 결정하는 방식이다. 즉, 그림 2와 같이 현재 프레임을 여러개의 기준 블록으로 나눈 다음, 각각의 기준 블록을 이전 프레임의 정해진 탐색 범위내에 있는 여러 개의 블록과 비교하여 가장 비슷한 매칭 블록을 찾아내어 기준 블록에 대한 매칭 블록의 상대 위치를 움직임 벡터로 정하는 기법이다.

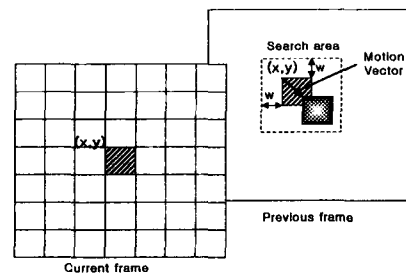


그림 2. 블록 매칭 알고리즘  
Figure 2. Block Matching Algorithm

블록 매칭 기법중 전역 탐색(Full-Search)은 기준 블록이 정해진 탐색 영역내의 모든 탐색점에 대하여 비교하는 방법이다. 이 방법은 오차는 적으나 모든 점에 대해 탐색을 하기 때문에 연산량이 많은 단점이 있다. 그러나 성능이 우수하고 하드웨어 구현이 간단하여 널리 쓰인다.

그래디언트 기법은 이미지 휘도가 변하지 않는다는 가정하에서 각각의 화소들의 실제 움직임을 추정하는 방법이다.[5][6]

블록 매칭 기법은 추정된 움직임 벡터가 실제 눈으로 느껴지는 움직임과 다를 때가 많으며 여러 개의 블록이 마치 모자이크처럼 보이는 블록킹 현상(blocking artifact)이 발생하는 단점이 있지만 연산량이 적고, 하드웨어 구현이 비교적 용이하며 성능이 우수하기 때문에 많이 쓰이고 있다.

## 3. RTP/RTCP 프로토콜

본 시스템에서는 실시간 영상 전송을 하기 위

해서 RTP/RTCP 프로토콜을 사용하였다.

RTP는 오디오/비디오 및 시뮬레이션 데이터와 같은 실시간 데이터를 멀티 캐스트 또는 유니캐스트 네트워크를 이용하여 전송하는 응용 서비스에 알맞는 단말-대-단말 네트워크 전송 기능을 제공한다.

이전의 RTP 라이브러리 설계와는 달리 본 개발에서는 트랜스포트 제어 메커니즘을 통합한다. 즉, 전송률 제어와 오류 제어를 RTP/RTCP 모듈에 통합한다. RTP는 1996년 IESG에 의해 RFC 1889로 채택되었고, 그 후 ITU-T 권고안 H.225.0의 일부가 되었다.

본 논문에서 구현한 RTP/RTCP 모듈에서 제공되는 기본 기능은 RTP 패킷 헤더 연산, RTP 페이로드 헤더 연산, RTCP 패킷 헤더 연산, 오류 제어와 RTP 계층 패킷 감시를 포함한다.

RTP/RTCP 모듈에는 RTP 흐름과 RTCP 패킷을 처리하고 관리하기 위한 요소들뿐만 아니라 다수의 부기(book-keeping) 요소들도 제공된다. 2대의 영상회의 단말이 사용할 미디어 유형에 동의하고 나면 그에 대응하는 RTP/RTCP 세션이 적당한 설정 정보(configuration information), 특히, 미디어의 페이로드 유형과 대역폭 요구를 가지고 개방된다. 이 설정 정보는 Bandwidth Manager와 Session Configuration Manager에 의해 계속 유지된다. 그리고 RTP/RTCP 모듈은 또한 대역폭이나 아니면 다른 시스템 자원이 불충분한 경우에 수락 제어(admission control)를 행한다[7].

RTP/RTCP에서의 동영상 패킷 전송은 카메라에서 비디오를 입력받아 코덱을 통해 압축되는데, 압축하면서 일정한 크기가 되면 이를 패킷화한다. 이에 비디오에서는 H.263이나 MPEG-2의 압축특성상 이전 프레임이 전송되지 않으면 다음 프레임을 복호화 할 수 없기 때문에 정확한 전송이 요구된다. 그래서 본 시스템에서는 비디오 패킷에 대해서는 시간내에 완전한 전송을 보장하도록 설계되었다. 일정 패킷을 캐스팅한 후에 확인신호를 일정시간동안 기다리게 된다. 대화참가자 확인신호가 모두 확인된 후에 다음 패킷을 전송하게 된다.

수신측에서는, 패킷 재조합 과정이 필요하므로 일정 크기의 버퍼를 사용하도록 하였다. 각 RTP 패킷 헤더에 타임스탬프와 시퀀스번호를 이용하여 패킷을 재조합하게 되는데, 일정시간내에 요구한 패킷이 수신되지 않으면 현재 수신된 패킷들을 가지고 디코딩을 수행하게 된다.[10]

### III. 디지털 감시 시스템

#### 1. 기존 시스템

기존의 아날로그 방식의 감시 시스템은 입력된 영상을 비디오 테이프에 녹화하여야 하며 이로 인하여 비디오 테이프를 자주 갈아 끼워야 하는 불편함과 화질이 선명하지 못하며 영상 처리 기능 면에서도 전진, 후진, 재생 기능 이외에는 제공하지 못한다는 단점을 가지고 있다. 최근에는 이러한 단점을 보완하기 위하여 디지털 방식의 감시 시스템(CCTV Monitoring System)이 개발되고 있다. 그러나 기존 시스템들은 대부분이 화질이 떨어지며 실시간 영상처리 속도를 유지하기 위해서는 고가의 전용 하드웨어를 사용하여 하는 문제점을 가지고 있다.

기존 디지털 감시 시스템의 가장 큰 단점 중에 하나는 데이터 저장 및 백업 문제이다. M-JPEG 또는 Wavelet 영상압축 기법을 이용한 감시 시스템에서는 데이터 저장을 위해 많은 양의 하드디스크 공간이 요구되며 이를 위한 하드디스크 및 백업 장비가 전체 시스템 구성에 많은 비용을 차지하고 있다. 이에 따라 감시 시스템을 개발할 때에는 영상 프레임 간의 화상정보 차이를 이용한 IPB 방식의 압축 기법을 통해 데이터 저장 및 전송시 압축률을 2배~5배 이상 높일 수 있는 영상 압축 기술인 MPEG 기법을 적용하고 있다.

기존의 디지털 감시 시스템에 사용되었던 영상 압축 기법들은 주로 H.263 또는 M-JPEG, MPEG-1의 변형을 사용해 왔고, 최근에는 MPEG-2를 사용하고 있다. 그러나 이러한 압축 기법들은 대부분 S/W적으로 구현되기보다는

H/W적으로 구현되기 때문에 시스템 구축, 유지 및 업그레이드 비용이 높아진다. 또한, 이러한 시스템들은 대부분 국제 표준을 정확하게 따르지 않고 있어 비호환성 문제를 지니고 있다.

따라서 본 연구에서는 높은 압축율과 실시간 인코딩 및 디코딩, 고품질의 영상 표현이 가능한 S/W 방식의 MPEG-2 디지털 감시 카메라를 개발하였으며, MPEG-2 국제 표준에 따라 정확하게 구현하여 압축된 영상정보의 호환성 문제를 해결하였다.

현재 국내의 디지털 감시 카메라 시장은 급속도로 커지고 있는데, 그 중에서도 많은 제품들을 내놓고 있는 3R(www.3R.co.kr)이나 디지털 영상 저장 장치(DVR)의 전문 업체인 성진 C&C(www.sjcnc.com)등이 선두를 달리고 있다. 많은 사업체들이 CCTV 시장에 뛰어들고 있다. 국내 외의 디지털 감시 카메라 시스템(CCTV) 시장에 관한 더 자세한 사항은 다음의 홈페이지(www.cctvbank.com)에 자세히 나와있다.

2. 제안된 감시 시스템

본 논문에서 제안한 시스템 구성도는 그림 3 과 같다.

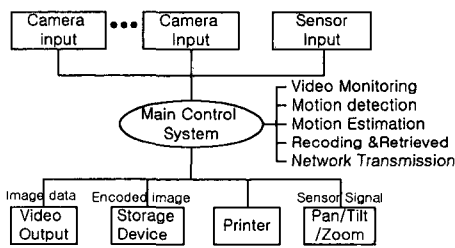


그림 6. 시스템 구성도  
Figure 3. System Configuration Diagram

시스템 구성에 대한 간략한 설명을 하면, 외부에서 최소 1대부터 최대 16대까지의 입력이 가능한 카메라 입력부, 센서의 신호를 입력을 받는 입력부와 움직임 검출, 비디오 감시, 물체 추적, 네트워크 전송, 움직임에 대한 녹화 및 복원 등을 제어하는 주 제어부, 그리고 녹화된 영상을

화면으로 출력하거나 부호화 된 영상의 저장, 프린트등을 담당하는 출력부로 나뉘어진다.

또한 본 시스템은 그림 4에서 나타낸 것처럼 IP 기반의 장거리 영상 전송과 원격지에서 감시할 수 있는 원격 모니터링 네트워크 기능을 가지고 있다. 그림 4는 인터넷이나 PSDN망을 통하여 카메라에서 입력받은 영상을 전송하는 구성도이다. 각 입력 카메라 영상의 1:1 뿐만 아니라 1:n도 가능하다.

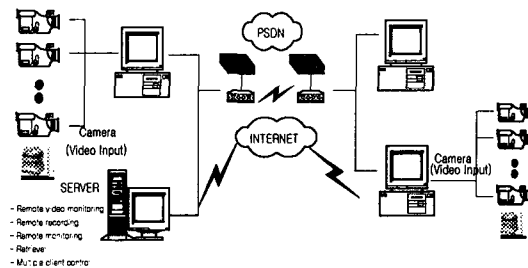


그림 4. 원격 모니터링 네트워크  
Figure 4. Remote monitoring Network

IV. 고속 움직임 추정과 물체 추출

1. 고속 움직임 추정

본 논문에서는 MPEG-2 인코딩을 위한 고속 움직임 추정 알고리즘을 제안한다. 본 알고리즘은 블록매칭 기법(block matching technique)과 그라디언트 기법(gradient technique)을 혼합하여 이용한다. 움직임 벡터의 추정에 대해, 블록 매칭 기법은 공간적으로 이웃하는 매크로블록들과 시간적으로 같은 매크로블록으로부터 초기 벡터를 결정하는데 이용된다, 그리고 나서 초기 벡터의 픽셀-귀환 세밀화 (pixel-recursive refinement)에 의해 최종 벡터를 얻기 위해 그라디언트 기법을 이용한다. 부가적으로, 그라디언트 기법이 연루될 때 저 지연 움직임 추정(low-latency motion estimation)을 가능케 하기 위해 검색 범위를 결정하기 위한 적응된 스키마(adaptive scheme)를 개발한다. 본 알고리즘이

전역탐색 블록매칭 기법(full search block matching)과 동등한 코딩 효율을 유지하는 동시에 빠른 속도도 가질 수 있다.

1.1 탐색 범위의 적응적 결정

그림 5는 움직임 추정을 위한 탐색 범위를 결정하는 적용된 스키마를 도시하고 있다. 적용된 스키마는 다음과 같은 4가지 단계로 구성된다:

- 탐색 범위의 적용된 결정
- 결정된 탐색 범위의 정규화
- 정규화된 탐색 범위의 스케일링
- 스케일된 탐색 범위의 표준화

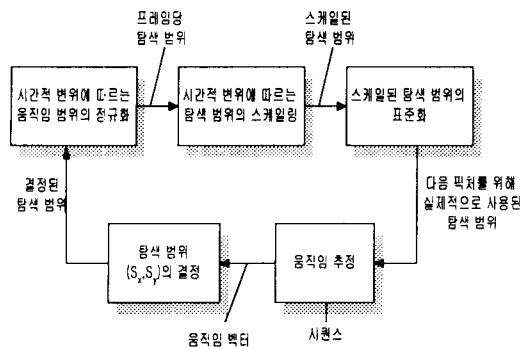


그림 8. 움직임 추정에서 탐색 범위의 적응적 결정

Figure 5. Adaptive determination of search range in motion estimation

수평 탐색 범위  $S_x$ 와 수직 탐색 영역  $S_y$ 는 독립적으로 결정된다. 절대적 움직임 벡터들이 이 결정에 이용된다.  $MB_{cnt}$ 를 픽처내의 총 매크로블록 수로 놓고,  $\{ (V_x^j, V_y^j) \mid j = 0, 1, 2, \dots, MB_{cnt}-1 \}$ 이 현재 픽처에서 적당치로 이용된 움직임 벡터라 할수 있으므로, 탐색 범위는 다음과 같이 정의된다.

$$\begin{cases} S_x = E(|V_x|) + 2 \cdot D(|V_x|) \\ S_y = E(|V_y|) + 2 \cdot D(|V_y|) \end{cases} \quad (1)$$

여기서,  $|\cdot|$ 는 절대 연산(absolute operation)을 의미하고,  $E(|V_x|)$ 와  $E(|V_y|)$ 는 수평과 수직 방향으로 절대적 움직임 벡터들의 평균 진폭들이다.

$$\begin{cases} E(|V_x|) = \frac{1}{MB_{cnt}} \sum_{j=0}^{MB_{cnt}-1} |V_x^j| \\ E(|V_y|) = \frac{1}{MB_{cnt}} \sum_{j=0}^{MB_{cnt}-1} |V_y^j| \end{cases} \quad (2)$$

곱셈연산을 피하기 위해, 편차(≠표준편차)  $D(|V_x|)$ 와  $D(|V_y|)$ 는 다음과 같이 정의된다.

$$\begin{cases} D(|V_x|) = \frac{1}{MV_{cnt}} \sum_{j=0}^{MB_{cnt}-1} \|V_x^j - E(|V_x|)\| \\ D(|V_y|) = \frac{1}{MV_{cnt}} \sum_{j=0}^{MB_{cnt}-1} \|V_y^j - E(|V_y|)\| \end{cases} \quad (3)$$

탐색 범위가 결정되어진 후에, 결정된 탐색 범위는 참조 프레임과 현재 프레임사이의 시간적인 거리를 이용하여 프레임당 탐색 범위(search range per frame)로 정규화된다. 그리고나서, 다음 픽처와 그 참조 픽처(들)사이의 시간적인 거리에 따르는 정규화된 탐색 범위를 스케일링함으로써 다음 픽처의 움직임 추정을 위한 탐색 범위가 얻어지게 된다.

마지막으로, 다음 픽처를 위한 스케일된 탐색 범위(순방향/역방향)이 표준화된다. 탐색 범위의 표준화는 동일한 f\_code들에서 최대 범위로 스케일된 탐색 범위를 반올림하는 것이다. S를 스케일된 탐색 범위(수평/수직/순방향/역방향)로 놓고, 탐색 범위는 표 1에서 보여진 것처럼 표준화된다.

표 1. 스케일 탐색 범위의 표준(in full-pixel units)

Table 1. Standard of Scale search range

range of S	std. S	f_code
[1:8]	8	1
[9:16]	16	2
[17:32]	32	3
[33:64]	64	4
[65:128]	128	5
[129:256]	256	6
[256:512]	512	7
[512:1024]	1024	8
[1024:2048]	2048	9

### 1.2 적응적 혼합 매칭 기법

그림 6은 세 개의 변형안을 고안함으로써 포맷 변환[9]을 위해 이용된 기법으로부터 얻어진 적응적 혼합 매칭 기법에 기반한 움직임 추정의 기본 알고리즘을 도시한 것이다. 포맷 변환[9]을 위한 움직임 추정에서는 블록 크기가 작은 반면, 여기서는 블록 크기가 MPEG-2 기술 때문에 16x16과 16x8로 제한된다. 블록들은 추정된 블록 벡터에서의 블록 스캔 순서 또한 [9]에서 이용된 것과는 다르다. 추가적으로, 추정된 탐색 범위를 저지연 코딩을 보장하기 위한 추정을 위해 고안되었다.

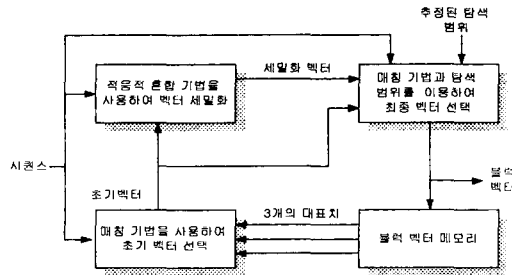


그림 6. 적응된 혼합 매칭 기법을 사용하는 움직임 추정의 블록도  
Figure 6. Block diagram of motion estimation using hybrid matching technique

대부분의 벡터 형태의 움직임 추정에 대해, 다음과 같은 3가지 과정이 포함된다.

- ° 매칭 기법을 이용하여 초기 벡터 선택
- ° 적응적 혼합 기법을 이용하여 벡터 세밀화
- ° 매칭 기법과 추정된 탐색 범위를 이용하여 최종 벡터 선택

무엇보다도, 픽처의 움직임 추정은 포맷 변환 [9]에서 사용된 순서와는 정확히 반대되는 미로 같은 순서에서 수행된다(그림 7). 움직임 추정에서 이 순서를 선택한 중요한 이유는 코딩 효율 때문이다. 실험적인 결과들로 이 순서가 매크로블록 코딩 순서(즉, top→bottom, left→right)보다 0.6 dB까지 성능이 뛰어나다는 것을 입증하였다. 이 스캔 순서의 결점은 인코딩시에 매크로블록 행의 지연을 야기시킨다는 것이다.

초기 벡터는 3개의 대표치들로부터 선택되어진다:

- ° 시간적 대표 벡터(temporal candidate vector)
- ° 수직 대표 벡터(vertical candidate vector)
- ° 수평 대표 벡터(horizontal candidate vector)

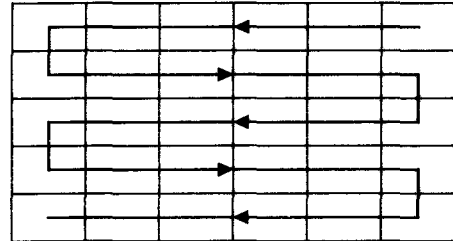


그림 10. 움직임 추정에 사용된 Meander scan  
Figure 7. Meander scan order used in motion estimation

각 대표 벡터에 대해 ADBD(the absolute displaced block difference)를 계산하는데 매칭 기법이 필요하게 된다.

초기 벡터로 가장 작은 ADBD 벡터를 선택한다. 만약 대표 벡터가 현재 매크로블록의 위치에 있을 때 참조 픽처의 바깥쪽을 가르키거나 이용할 수 없다면, 이 벡터는 선택하는데 고려되지 않는다.

두 번째 단계는 그래디언트 기법을 이용하여 초기 벡터를 세밀화하는 것이다. 여기에 픽셀 귀환[3]이 적용된다. 움직임 보상된 시간적 그래디언트와 픽셀 위치에서 공간적 그래디언트(수직/수평)가 계산된다. 그러면 초기 벡터가 픽셀×픽셀로 세밀화되고, 마지막으로 세밀화된 벡터는 가장 작은 DPD의 하나로서 선택된다.

마지막 단계에서, ADBD가 세밀화된 벡터에 대해 계산된다. 만약 세밀화된 벡터나 초기 벡터 모두가 추정된 탐색 범위내에 있다면, 더 작은 ADBD 벡터가 최종 벡터로 선택된다. 만약 세밀화된 벡터나 초기 벡터 어느 하나라도 탐색 범위 밖에 있으면, 탐색 범위내의 벡터가 선택된다. 바꾸어 말하면, 이들 벡터 모두가 탐색 범위를 초과한다면, 더 작은 ADBD 벡터가 단순히 탐색

범위내로 클리핑 된다는 것이다

MPEG-2 비디오 인코딩을 위한 빠른 움직임 추정 알고리즘은 적용된 스키마와 위에서 기술된 기본적인 움직임 추정 알고리즘의 반복된 사용으로 구성된다. 적용된 스키마는 탐색 범위를 결정하기위해 사용되고, 기본적인 움직임 추정 알고리즘은 매크로블록×매크로블록으로 벡터들을 추정하기 위해 사용된다.

2. 움직이는 물체 추출

본 논문에서, 부호화된 데이터 영역에서 MPEG-2 비트 스트림으로부터 효과적인 움직이는 물체 검출 알고리즘을 제안한다. 본 알고리즘에서, 부호화된 비트 스트림에서 직접적으로 추출될 수 있는 움직임 정보 같은 부호화된 정보만을 이용한다. MPEG-2에서 직접 얻은 움직임 정보들은 조명이나 하드웨어적인 특성 때문에 노이즈를 가지게 된다. 이러한 노이즈로 인해서 움직임 정보를 물체 추출에 직접 이용하는 것은 오류를 발생시킬수 있다.

이러한 문제를 해결하기 위해서 본 논문에서는 노이즈 제거를 위해 피크 노이즈에 강한 중간값(Median)[2] 연산자를 사용하였다.

움직임 정보가 물체의 움직임을 나타낼지라도, 모든 움직임 정보가 항상 실질적인 움직임을 나타내지는 않는다. 그 이유는 움직임 추정이 물체의 움직임 관점에서 수행되는 것이 아니라 부호화 효율성 관점에서 수행되기 때문이다. 부호화된 데이터로부터 물체의 실질적인 움직임을 얻기 위해서, 본 논문에서는 부호화된 데이터 영역에서의 유용한 정보를 이용하여 시/공간적 움직임 연속성 같은 물체의 움직임 특징을 연구하였다.

2.1. 움직임 정보가 존재할 경우

(1) 공간적인 상관성

일반적으로, 동일하게 움직이는 영상 내의 움직이는 물체는 공간적으로 높은 상관도를 가지고 있다. 그림 8의 블록 MB(i,j)는 움직임에 있어서, 인접한 블록에 높은 상관성을 지니고 있다.

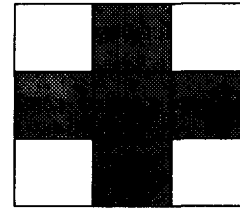


그림 8. 움직임의 공간적인 상관성  
Figure 8. Spatial correlation of Motion

공간적으로 이웃하는 8개의 매크로블록들을 조사하여 적어도 그 벡터들중의 하나가 비슷한 움직임 벡터를 갖는다면 타겟 매크로블록이 움직이는 물체 후보 매크로블록에 포함되는지를 결정한다. 유사도 측정법처럼, 내삽적(inner product)을 취함으로써 움직임 벡터들의 각도 차분을 이용한다. 다음 수식이 이것을 나타낸다. 즉,

$MB_n(i,j)$  is moving object candidate

$$\text{if } \frac{MV_n(i,j) \cdot MV_n(i+k,j+l)}{|MV_n(i,j)||MV_n(i+k,j+l)|} > Thsc,$$

$$k, l = \{-1, 0, 1\}, (k, l) \neq (0, 0) \quad (4)$$

여기서,  $MB_n(i,j)$ 는 n번째 프레임에서 (i,j)번째 매크로블록을 의미한다.  $MV_n(i,j)$ 는 매크로블록 (i,j)의 움직임 벡터이다. 그리고 Thsc는 공간적 상관관계 검출을 위한 임계값이다. 도트는 내삽적 연산을 의미하고 ||는 노름 연산(norm operation)을 의미한다.

(2) 시간적인 상관성

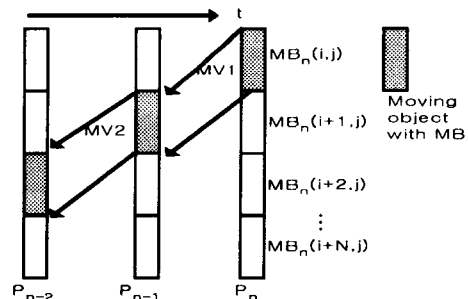


그림 9. 움직임의 시간적인 상관성  
Figure 9. Temporal correlation of motion



만약 표적 매크로 블록이 움직이는 물체를 포함한다면, 움직임 보상을 위해서 사용한 참조 프레임에서의 지역은 그림 9에서 보여준 것처럼 현재의 매크로 블록에서 매우 비슷한 움직임 특성들을 가지게 될 것이다.

움직이는 물체 움직임을 가지는 현재의 매크로 블록 MBn(i,j)는 MV1을 사용하는 Pn-1 현재 프레임의 지역에서 보상된다.

시간적인 상호관계를 고려하면, 움직이는 물체를 포함하는 지역의 대부분을 덮는 그림에서 매크로 블록 MBn(i+1,j)는 상당히 가능성이 있는 일이다.

그때, 만약 움직이는 물체가 지역에 포함되어 있다면, 매크로 블록 MBn-1(i+1, j) 움직임은 MV1에서 매우 비슷한 벡터를 가지는 MV2움직임 벡터를 사용하여 Pn-2프레임안에서 상응하는 지역을 보상한다. 이러한 시간적인 상관성을 사용함으로써, 만약 이러한 움직임 벡터들의 각도 차이가 작다면, 우리는 움직이는 물체 후보자들에 속하는 매크로 블록을 결정한다. 즉,

MBn(i, j) is moving object candidate

$$\text{if } \frac{MV_n(i, j) \cdot MV_{n-m}(p, q)}{|MV_n(i, j)| |MV_{n-m}(p, q)|} > Thtc \quad (5)$$

이다.

여기에서, Thtc는 시간적인 상관성 검출을 위한 임계값이고, MVn-m(p,q)는 n-m번째 프레임의 (p,q) 매크로 블록에서 움직임 벡터이다.

여기서, n-m번째 프레임은 n번째 프레임의 참조 프레임이다.

MBn-m(p,q)는 MBn(i,j)의 움직임 보상된 지역에서 가장 관계되는 매크로 블록이다.

매크로 블록의 움직임이 참조 프레임에서 한 지역을 보상할 때, 4개의 매크로 블록에 속하는 지역에서 참조할 것이다.

관련이 가장 많이 되는 매크로블록 MBn-m(p,q)에서 움직임 보상 지역의 25%이상 덮여진 매크로 블록을 결정한다.

후향(backward) 움직임 보상 경우에 관해서는, MVn(i,j) 대신에 뒤집혀진(reversed) 움직임 벡터-MVn(i,j)을 사용한다.

이러한 방법은 B-픽처에서 양방향적으로 움직임 보상되는 경우에 쉽게 전개되어질수 있다.

이러한 경우에서, 타겟 매크로 블록은 미래 프레임과 과거 프레임을 사용하는 두 개의 움직임 벡터들을 가진다.

적은 프레임 거리에서 선형적으로 움직임 모델 이상을 적용함으로써, 여기서 연속되는 두 개의 프레임들에서 두 개의 움직임 벡터들 대신에 전향과 후향 움직임 벡터들을 사용한다, 만약 두 개의 벡터들이 거의 반대 방향이라면 움직이는 물체에 속하는 매크로 블록을 결정한다.

이러한 관계는 아래의 방정식으로써 표현되어질 수 있다.

MBn(i, j) is moving object candidate

$$\text{if } \frac{MBF_n(i, j) \cdot (-MVB_n(i, j))}{|MVF_n(i, j)| |MVB_n(i, j)|} > Th tcb \quad (6)$$

여기서, MVFn(i,j)와 MVBn(i,j)는 각각 전향과 후향 움직임 벡터들이다.

## 2.2 움직임 정보를 포함하지 않을 경우

MPEG에서 움직임 정보를 포함하지 않는 매크로 블록의 코딩 방법은 다음 세 가지가 있다.

No-MC No-Coded(skip MB), No-MC Coded MB 과 intra coded MB이 있다.[1][2]

처음 두 가지 경우에서, No-MC(no motion compensation) 모드들은 움직임 벡터가 (0,0)인 이후부터 선택되어진다. 그러므로, 실제 움직임은 이러한 모드들에 포함되지 않는 것으로 가정되어질 수 있다.

인트라 부호화 된 매크로 블록들에서, 두가지 경우에서 더 세분화되어진다.

하나는 인트라 코딩인데 P-와 B-픽처에서 효율적인 코딩에 의해서 선택되어진다. 그리고 다른 하나는 매크로 블록은 강제로 I-픽처에서 인트라 부호화되어진다는 것이다.

그림 10은 움직이는 물체를 검출하는 전반적인 흐름도를 보여준다.

그림 10은 MPEG 비디오 스트림을 인트라 부호화된 MB으로 분리하여 인트라 부호화된 MB 이면 추출된 텍스처 정보가 I-픽처와 그렇지 않

은 경우로 나누어서 움직임 벡터(MV) 객체 후보자들을 선택하고, 그렇지 않으면 흐름도에 나와있는 방법을 시행한다. 그후에 중간값 필터를 사용하여 노이즈를 제거하고, 이전의 텍스처 정보를 사용하여 플랫폼 지역을 제거하고 고립된 매크로 블록들을 포함/제거 하는 단계를 거쳐서 움직이는 물체 매크로 블록을 검출한다.

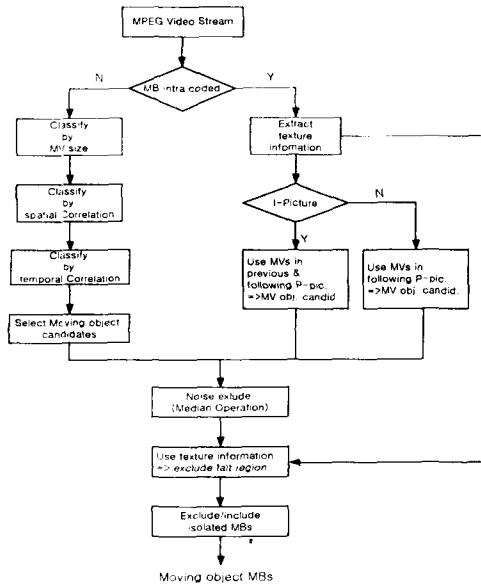


그림 10. 움직이는 물체 검출 알고리즘  
Figure 10. Moving object detection algorithm

### V. 실험 및 분석

본 연구에서는 먼저 MPEG-2 인코더 및 디코더를 Pentium III PC에서 Visual C++ 6.0을 사용하여 소프트웨어로 구현하였으며, OS는 마이크로 소프트의 윈도우즈 환경에서 구현하였다.

압축된 영상 정보의 실시간 전송을 위해 RTP/RTCP 기법을 적용하였다.

본 논문에서 제안한 빠른 움직임 추정 알고리즘의 성능 평가를 위해서 FS(full search), DS(diamond search), TSS(three step search) 알고리즘과의 PSNR을 비교하였다.[10][11][12]

탐색 범위를 결정하는 적응된 스키마는 움직

임 추정 알고리즘 모두에서 이용된다.

실험에서 이용된 주된 코딩 파라미터들은: N=12, M=3, bit-rate 3 MBit/s이다. 움직임 벡터들은 원래의 시퀀스(original sequence)들 위에서 추정된다.

표 2. 프레임구조에서 PSNR[dB] 비교표  
Table 2. Comparison of PSNR in frame structure

	Basket	Flower	Mobile
FS	28.96	30.57	28.42
TSS	27.55	28.76	27.01
DS	28.17	29.11	27.54
Fast	28.48	29.94	28.03

표 3. 필드구조에서 PSNR[dB] 비교표  
Table 3. Comparison of PSNR in field structure

	Basket	Flower	Mobile
FS	29.56	31.30	29.08
TSS	28.34	28.97	27.91
DS	28.97	29.44	28.27
Fast	29.05	30.81	28.83

실험 결과인 표2와 3에서 보는 것처럼 다른 알고리즘과 비교하여, 본 논문에서 제안하는 알고리즘의 손실이 상대적으로 적은 것을 알 수가 있다.

그림 11과 그림 12는 위에서 언급한 알고리즘을 가지고 실내에서의 움직임이 검출되었을 경우에 움직임이 검출되었음을 보여준다.



그림 14. 원 영상  
Figure 11. Original Image

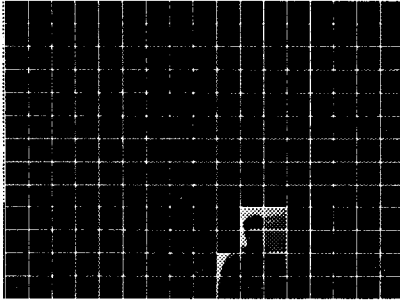


그림 15. 움직임 검출된 영상  
Figure 12. Detected Moving object

다음은 실외에서의 움직임 검출에 관한 실험 결과이다.



그림 16. 원 영상  
Figure 13. Original Image



그림 17. 움직임 검출 영역  
Figure 14. Moving detection area

제안된 고속 움직임 추정 알고리즘은 MPEG-2 비디오 인코딩시에 높은 속도의 움직임 추정이 가능하게 한다. 반면 여전히 높은 코

딩 효율과 작은 지연을 가능토록하는 문제는 남아있다. 그러므로, 이 알고리즘은 거의 실시간적인 소프트웨어 MPEG-2 인코딩의 솔루션과 단일칩 하드웨어 구현과 같은 어플리케이션에 크게 도움이 된다.

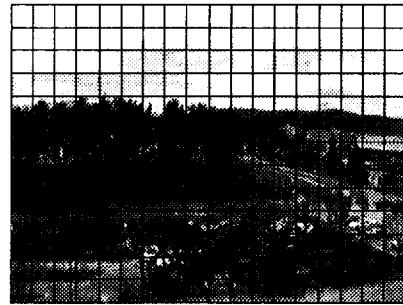


그림 18. 움직임 검출된 부분  
Figure 15. Detected moving object

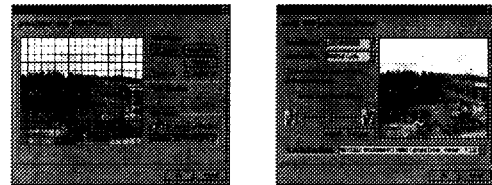


그림 16. 시스템의 Setup 화면  
Figure 16. Setup display of System

본 논문의 고속 움직임 검출 알고리즘과 고속의 움직임 추정 알고리즘을 이용하여 약 30frame/s의 전송속도를 가지는 보다 실시간적인 S/W 인코더의 구현을 가능하게 하였다. 또한 본 논문에서 제안한 시스템은 기존의 시스템들과는 달리 국제 표준인 MPEG-2에 근거하여 최대한 따르려고 노력하였다.

## VI. 결론 및 향후 연구방향

본 논문에서 제안된 고속 움직임 추정 알고리즘은 MPEG-2 비디오 인코딩시에 높은 속도의

움직임 추정을 가능케 한다. 반면 여전히 높은 코딩 효율성과 저 지연을 가능토록 하는 문제가 남아있다. 또한 제안된 알고리즘은 적당한 변경을 통해 MPEG-4 object-based 비디오 인코딩에 이용될 수 있을 것이다. 부호화 된 MPEG 데이터를 사용하여 움직이는 물체를 추출하는 알고리즘을 서술하였다. 부호화 된 데이터 영역에서 물체의 움직임을 특성화함으로써, 단지 움직이는 물체들은 움직임 정보와 DCT 계수들로부터 매크로 블록 레벨을 검출할 수 있다. 모든 필요한 정보가 부호화 된 데이터로부터 직접적으로 추출되어지기 때문에, 매우 빠른 동작으로 실현되어진 것을 알 수 있다. 하나의 움직임을 가지는 물체의 검출에 관해서만 논의하였지만, 이러한 것들은 원격 감시 응용에는 충분할 것이다. 그러나 TV나 영화 같은 실제적인 비디오 시퀀스들에서 제안하는 비디오 인덱싱을 위해서 사용되어 질 수 있게 하기 위해서 제안될 알고리즘을 좀 더 연구하여 여러 개의 움직이는 물체들의 검출을 위해 확장하여야 할 것이다.

본 논문에서 제안한 움직임 검출 감시 시스템은 뛰어난 압축률을 보이는 MPEG-2를 이용하여 모든 코덱을 100% 소프트웨어로 구현하였다. 하지만 윈도우 운영체제하에서 운영되어 안정성 면에서 조금 미흡한 면이 없지 않다. 이를 위해 리눅스를 기반으로 한 내장된(embedded) 시스템을 설계할 필요성이 있다. 또한, 단순하게 저장만을 위한 시스템이 아닌 사물을 인식하여 판단할 수 있는 적용된 영상인식을 연구하여 본 시스템에 적용시켜, 무인감시시스템의 목적인 사람이 필요 없는 시스템에 조금 더 가깝게 다가갈 수 있도록 하겠다.

향후에는, 본 시스템을 좀더 개발하여 감시되고 있는 수십 대 이상의 카메라 중에서 어느 한 카메라에 움직임 검출이 나타나면 움직임 검출이 된 카메라의 위치가 지도에 표시가 되는 시스템을 연구 개발할 예정이다.

### 참고문헌

[1] ISO/IEC 13818-2, ITU-T Rec. H.262: 1995,

Information technology - Generic coding of moving pictures and associated audio information - Part2: Video

[2] A.Murat Tekalp, "Digital Video Processing", Prentice Hall PTR, University of Rochester, pp. 95-108, 1995

[3] Joan L.Michell, William B.Pennebaker, Chad E.Fogg, Dider J.LeGell, "MPEG Video Compression Standard", Champman & Hall, New York, pp.237-262, 1997

[4] ISO/IEC 13818-1,2,3, ITU-T Rec. H.262: 1995, Information technology - Generic coding of moving pictures and associated audio information - Part2: Video

[5] F. Dufaux, "Motion Estimation Techniques for Digital TV: A Review and a New Contribution", Proc. Of the IEEE, Vol. 83, No.6, pp.858-876, June 1995.

[6] L.K. Liu and E.Feig, "A block-based gradient descent search algorithm for block motion estimation in video coding," IEEE Trans. Circuits Syst. Video Technol., vol.6 pp.419-423, Aug 1996

[7] Network Working Group, RFC 1890, RTP Profile for Audio and Video Conferences with Minimal Control, 1996

[8] J. Ohm, E. Guettner, "Concepts of Conventional Format Conversion", MINT/AG2 /HHI/032.a4.

[9] ISO/IEC 13818-1,2,3, ITU-T Rec. H.262: 1995, Information technology - Generic coding of moving pictures and associated audio information - Part2: Video

[10] F. Dufaux, "Motion Estimation Techniques for Digital TV: A Review and a New Contribution", Proc. Of the IEEE, Vol. 83, No.6, pp.858-876, June 1995.

[11] L.K. Liu and E.Feig, "A block-based gradient descent search algorithm for block motion estimation in video coding," IEEE Trans. Circuits Syst. Video Technol., vol.6 pp.419-423, Aug 1996

[12] S. Zhu and K.-K.Ma, "A new diamond search algorithm for fast block matching motion estimation," IEEE Trans. Image Process., vol.9, pp.287-290, Feb. 2000.



김선우(Sun-woo Kim)  
현재 군산대학교 전자정보공  
학부 대학원 석사과정

※ 관심분야 : 영상 처리, JPEG-2000, 웨이블릿,  
워터마킹



최연성(Yeon-Sung Choi)  
현재 군산대학교 전자정보공  
학부 부교수

※ 관심분야 : 영상 처리, 멀티미디어 시스템, 신호  
처리 시스템



박한엽(Han-Yeop Park)  
현재 군산대학교 대학원 정  
보통신공학과(박사과정)

※ 관심분야 : 영상 통신, 비디오 코딩