

논문 2006-43SP-4-11

파노라믹 영상 모델에 근거한 파라메트릭 비디오 압축

(Parametric Video Compression Based on Panoramic Image Modeling)

심 동 규*

(Donggyu Sim)

요 약

본 논문에서는 파노라믹 영상 모델링에 근거한 비디오 압축 전송 방법을 제안한다. 제안한 방법은 회전 카메라에 입력되는 영상에서 배경 영상과 움직이는 물체로 분리하고, 차영상을 추출하여 압축/전송하는 방법을 사용한다. 제안한 비디오 압축 시스템은 초기화 과정에서 전송된 파노라믹 영상으로부터 배경영상을 합성할 수 있도록 파라미터 만을 전송하게 된다. 본 논문에서는 정확한 배경 합성을 위한 정확한 카메라 모델링 기반 파노라믹 영상 합성법을 제시하며, 이를 바탕으로 비디오 압축에 응용하는 방법을 제안하였다. 제안한 비디오 압축방법에 의하여 기존의 JPEG-2000이나 MPEG-4 비디오 압축 방법에 비하여 PSNR 관점에서 2~4dB 효율적임을 보였다.

Abstract

In this paper, a low bitrate video coding method based on new panoramic modeling is proposed for panning cameras. An input video frame from a panning camera is decomposed into a background image, rectangular moving object regions, and a residual image. In coding the background, we employ a panoramic model that can account for several image formation processes, such as perspective projection, lens distortion, vignetting and illumination effects. Moving objects are detected, and their minimum bounding rectangular regions are coded with a JPEG-2000 coder. We have evaluated the effectiveness of the proposed algorithm with several indoor and outdoor sequences and found that the PSNR is improved by 1.3~4.4 dB compared to that of JPEG-2000.

Keywords : 보안 카메라, 비디오 코딩, 패닝 카메라, JPEG-2000, MPEG

I. 서 론

현대 사회에서 감시카메라는 사건/사고의 간접적 예방 및 증거확보를 위하여 폭넓게 사용되고 있다^[1]. 최근 들어서는 디지털 영상 신호처리 기술의 발전으로 보다 저렴하고 소형화된 디지털 방식의 감시카메라 시스템이 상용화되고 있다^{[1]-[6]}. 이러한 디지털 비디오 압축 기술에 기반을 둔 방식은 기존의 아날로그 테이프 방식에 비하여 관리 비용이 적게 들며, 화질의 저하가 없는 장점을 가지고 있다. 그러나, 아직도 장시간의 비디오 압

축시 대용량 하드 디스크를 필요로 하는 문제점을 가지고 있다. 현재 폭 넓게 사용되고 있는 회전 카메라의 경우, 같은 배경을 반복적으로 취득하고, 이를 반복적으로 압축/저장 혹은 전송하고 있다. 이러한 반복적으로 나타날 배경을 파노라믹 영상으로 표현하고, 이를 사전에 전송하여, 배경부분을 반복적으로 압축/저장/전송하는 비용을 줄이는 시스템을 제안하였다.

감시 카메라를 위한 이상적인 비디오 압축 방법은 고 압축율, 여러 강건성, 신뢰성과 검색의 용이성이 요구된다. 현재 연구 개발되고 있는 DVR (Digital video recorder)는 동영상 압축 표준인 MPEG-4 part 2^{[6][7]} 혹은 JPEG-2000^{[8][9]}를 사용하고 있다. 최근 들어 H.264/AVC를 기반으로 하는 DVR에 대한 연구도 진행되고 있다. 이러한 동영상 압축기법에 근거한 감시카메라 시스템은 고 압축율을 이룰 수 있으나, 검색 등이 용이하

* 정회원, 광운대학교 컴퓨터공학과
(Kwangwoon University)

※ 이 논문은 2005년 정부의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임.

(KRF-2005-003-D00359)

접수일자: 2006년3월30일, 수정완료일: 2006년6월2일

지 않고 계산량이 많아 높은 비용을 갖는 단점을 가지고 있다. 이러한 이유로 실제로 상용화되고 있는 많은 DVR 장치들은 움직임 검출 장치와 결합된 정지영상 압축 기법이 사용되고 있다. 또한 동영상 압축 기법은 입력 영상의 내용과 프레임 레이트와 같은 코딩 변수들에 의하여 성능 변화가 큰 단점을 가지고 있다. 실제로 보안 목적 카메라의 경우 빠른 움직임이 있는 부분이 가장 중요한 영상이 되나, 실제로 동영상 압축기법의 경우 이러한 영상에서 화질의 저하라는 큰 단점을 가지고 있다. 또한 감시 카메라의 경우 여러 대의 카메라를 동시에 사용해야 하므로 낮은 프레임 레이트가 보편적이나, 동영상 압축 기술의 경우 낮은 프레임 레이트에서 큰 성능 저하를 가지게 되는 단점을 가지고 있다. 한편, 회전 카메라를 위한 비디오 압축 기법으로 파노라믹 영상 합성^{[10]-[12]}에 근거한 비디오 압축 방법이 제안되기도 하였다. 이러한 연구는 배경 영상 파라미터 전송과 합성 배경영상과 원영상의 차영상을 압축하는 방

법으로 비디오 압축을 이룬다^[13]. 이러한 방법은 배경영상과 원영상 간의 차 영상이 비디오 압축기법에 적합하지 않아 압축이 크게 향상되지 않는 단점을 가지고 있다.

본 논문에서는 하나의 카메라로 넓은 지역을 감시할 수 있는 회전카메라의 경우 반복되는 배경을 효과적으로 표현/압축하는 방법을 제안하고, 이를 바탕으로 입력되는 영상에서 배경과 움직이는 물체를 분리하여 움직이는 물체만을 압축 전송하는 시스템을 제안하였다. 여기서 정확한 배경 생성을 위하여 기존의 방법에 비하여 효과적인 카메라 모델링 기법을 사용하였으며, 배경과 움직이는 물체를 분리하는 방법과 전체 비디오 압축 시스템을 제안하였다. 본 방법은 움직이는 물체를 분리하고, 시간적 중복성을 사용하지 않는 JPEG-2000 방법을 사용함으로써, 효과적인 검색과 영상화질을 유지하였다. 특히, 증거확보에 있어서 움직이는 물체의 중요성을 볼 때, 움직임 물체 중심의 감시카메라가 필

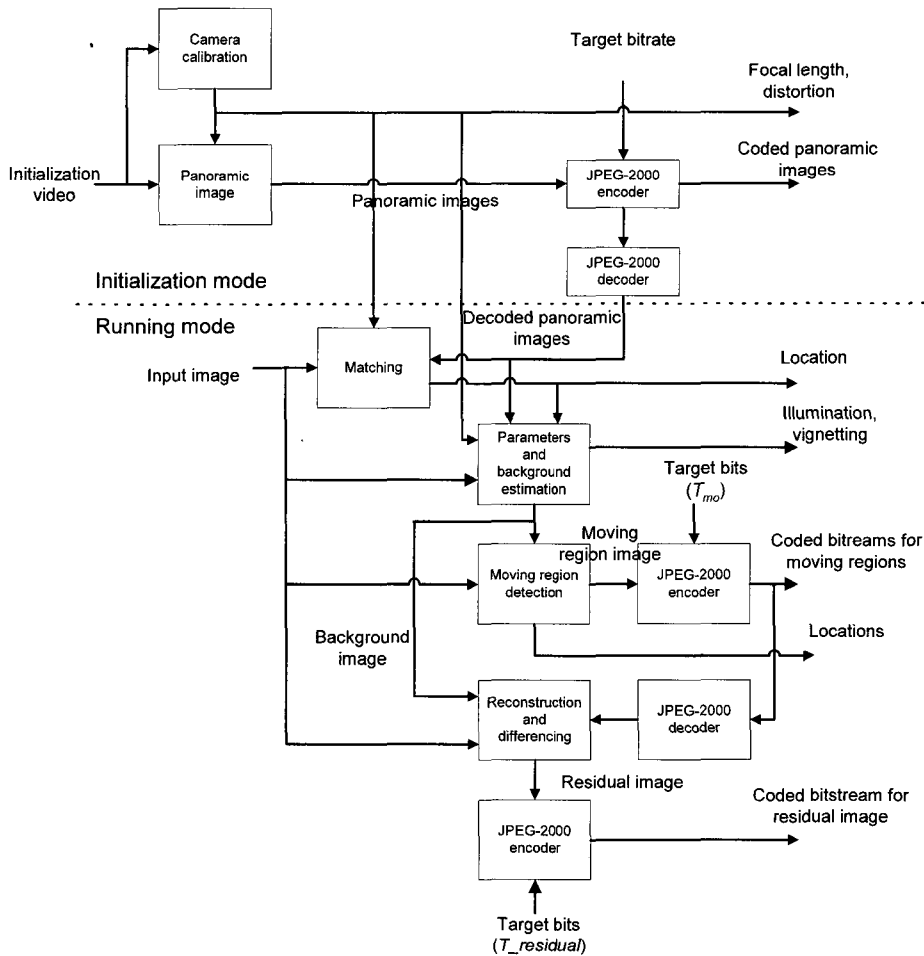


그림 1. 제안한 비디오 압축 기법의 블록도

Fig. 1. Overall block diagram of the proposed video coding algorithm.

수적이라 할 것이다. 본 논문에서는 제안한 방법의 효율성을 보이기 위하여 기존의 JPEG-2000방법과 MPEG-4 비디오 압축방법의 PSNR을 비교함으로써, 제안한 방법이 기존의 기술들에 비하여 2-4dB 정도 효과적임을 보였다.

II. 제안한 파노라믹 영상 모델링에 근거한 비디오 압축

그림 1은 제안한 파노라믹 영상 모델링 기반 비디오 압축 블록도이다. 제안한 비디오 압축 시스템은 파노라믹 영상 합성부, 배경 영상 복원부와 움직임 물체 검출부로 구성되어 있다. 그림 2는 그림 1의 부호화에 대한 복호화 방법을 나타낸다. 제안한 파노라믹 영상 기반 비디오 압축은 그림 1에서와 같이 초기화 시, 회전 카메라에 의하여 나타날 배경을 완전히 포함할 수 있는 배경영상을 압축 전송한다. 이 배경영상을 전송할 때 카메라 초점거리와 렌즈 왜곡 파라미터를 함께 전송하게 된다. 이러한 초기화 모드에서 실제 동작모드로 시스템이 전환된 후, 입력 비디오에 해당하는 배경영상 부분을 전송하지 않고, 몇 개의 파라미터를 전송하게 된다. 복호화 쪽에서는 이전에 전송된 전체 배경영상과 카메라 파라미터를 이용하여, 현재 입력되고 있는 영상

에 해당하는 배경영상을 합성 재생한다. 이 때, 초점거리, 비그네이팅 파라미터, 렌즈왜곡 및 조명 파라미터 등을 사용하여 입력 영상을 정확히 모델링할 수 있게 된다. 그림 1에서 실제 입력 영상이 배경만이 존재하는 것이 아니라, 움직임 물체가 포함되어 있을 경우, 생성된 배경영상을 이용하여 움직이는 물체를 추출하게 된다. 이 움직이는 물체를 포함하는 최소 직사각형 영역에 JPEG-2000 압축기법을 이용하여 움직임 물체를 압축하고, 이를 통하여 복원된 영상을 이용하여 차영상을 생성하고, 차영상을 다시 JPEG-2000으로 코딩하여 전송하게 된다.

1. 카메라 모델링 및 보정

카메라로부터 얻은 영상은 카메라의 내부 및 외부 파라미터에 따라 그 영상의 특성이 크게 달라진다. 실제 카메라 기반 영상 분석에서 카메라의 정확한 모델링은 전체 시스템의 성능을 좌우하는 큰 요소가 되고 있다 [13]. 본 논문에서는 그림 3과 같이 카메라 영상 모델링을 위하여 원근 투시법, 렌즈 왜곡, 조명 변화 및 비그네이팅 파라미터를 사용하였다. 렌즈 왜곡 보정은 저가의 카메라에서는 필수적인 요소로 고려되어야 한다.

큰 렌즈 왜곡이 존재할 경우, 3차원 공간에서 직선이 영상에서 구부러진 곡선으로 나타나게 된다. 이러한 렌

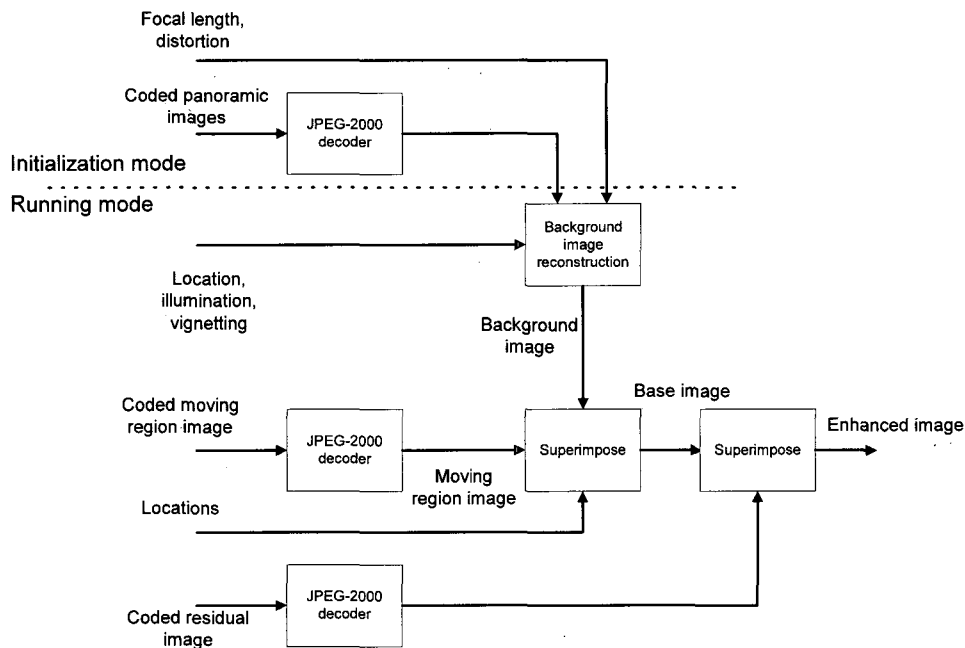


그림 2. 제안한 비디오 복호화기 블록도

Fig. 2. Overall block diagram of the proposed decoding system.

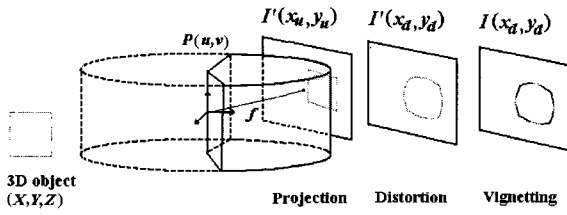


그림 3. 카메라 및 실린더 좌표 투영 모델링
Fig. 3. Image formation process model.

즈 왜곡을 다음의 식으로 모델링 한다^{[14]-[16]}.

$$\begin{aligned} x_u &= x_d \{1 + K_1(x_d^2 + y_d^2) + K_2(x_d^2 + y_d^2)^2\} \\ y_u &= y_d \{1 + K_1(x_d^2 + y_d^2) + K_2(x_d^2 + y_d^2)^2\} \end{aligned} \quad (1)$$

여기서 (x_d, y_d) 과 (x_u, y_u) 는 왜곡 영상 좌표와 무 왜곡 영상 좌표를 나타낸다. 본 논문의 영상 압축 시스템은 초기화 단계에서 K_1 과 K_2 파라미터가 예측된다. 여기에 조명 변화와 함께 비그네이팅 효과를 함께 모델링하기 위하여, 다음의 수학적식을 사용한다^[14].

$$I'(x_d, y_d) = \frac{L^0 I(x_d, y_d) + L^1}{V \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{x_d^2 + y_d^2}}{f} \right) \right) + 1 - V} \quad (2)$$

여기서 L^1 와 L^0 는 조명 변화 파라미터를 나타내고, V 는 비그네이팅 파라미터를 나타낸다. 는 영상의 화소 간의 실제 간격을 나타내며, f 는 카메라의 초점거리를 나타낸다. 조명 변화는 주변 환경의 밝기 조건의 변화에 따라 결정되며, 비그네이팅 효과는 영상의 주변으로 갈수록 렌즈 시스템에 의하여 입사 빛이 차단되어 나타내는 현상으로, 영상의 주변으로 갈수록 영상이 어두워 보이는 현상을 말한다.

2. 파노라믹 영상 합성

그림 4는 카메라의 초점에서 회전하는 카메라에 의하여 취득된 다수의 영상으로부터 실린더 평면에 파노

라믹 배경영상을 투영하는 구조와 블록도를 나타낸다. 본 투영을 위하여, 실린더 평면의 좌표 (u, v) 와 무 왜곡 평면의 좌표 (x_u, y_u) 와는 다음과 같은 관계를 가지게 된다.

$$x_u = \tan\left(\frac{u}{f}\right), \quad y_u = \frac{v}{\cos\left(\frac{u}{f}\right)} \quad (3)$$

여기서 (u, v) 실린더 좌표와 연관된 무 왜곡 좌표 (x_u, y_u) 는 식 (1)과 (3)을 통하여 유도할 수 있다. 식 (1)이 5차 다항식이어서, 고차 다항식의 해를 구해야 하는 문제가 발생한다. 그러나, 렌즈가 배럴 왜곡이고, K_2 가 0이라고 가정한다면, 한 개의 실수 해를 양의 함수 (explicit function)로 나타낼 수 있게 된다. 이에 대한 실수해 x_d 는 다음과 같이 나타낸다.

$$\begin{aligned} Q_x &= -\frac{x_u^2}{3K_1(x_u^2 + y_u^2)}, \quad R_x = \frac{x_u^3}{2K_1(x_u^2 + y_u^2)} \\ A_x &= \begin{cases} -(R_x + \sqrt{R_x^2 - Q_x^3}), & R_x > 0 \\ (-R_x + \sqrt{R_x^2 - Q_x^3}), & \text{otherwise} \end{cases} \\ x_d &= \begin{cases} A_x + \frac{Q_x}{A_x}, & A_x \neq 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (4)$$

y_d 는 x_d 가 구해진 것과 같은 방법으로 구할 수 있을 것이다. 파노라믹 영상 합성을 위하여, 역 워핑 기법을 사용하여 실린더 평면 영상 $P_m(u, v)$ 를 생성한다.

그림 4에서와 같이, 하나의 파노라믹 영상을 구성하기 위하여, 다수의 실린더 평면 영상에서 특징점을 추출하고, 이들의 대응점을 찾음으로써 하나의 파노라믹

영상을 구성하게 된다. 여기서 특징점을 추출하기 위하여 연속되는 두 영상이 겹쳐지는 부분에서 최대 분산 값을 갖는 위치를 추출한다. 이렇게 추출된 특징점에서 그림 5와 같이 블록 기반 정합을 통하여 대응점을 추출한다. 연속된 두 영상을 합성하기 위하여 비그네이팅 파라미터 (V_p)와 조명 파라미터는 다음의 오차식을 최

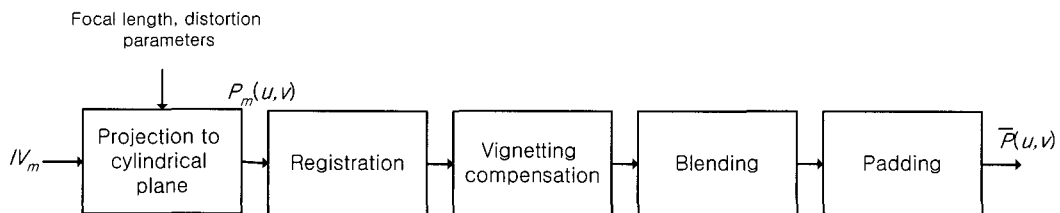


그림 4. 파노라믹 영상 합성 블록도
Fig. 4. Block diagram for panoramic image generation.

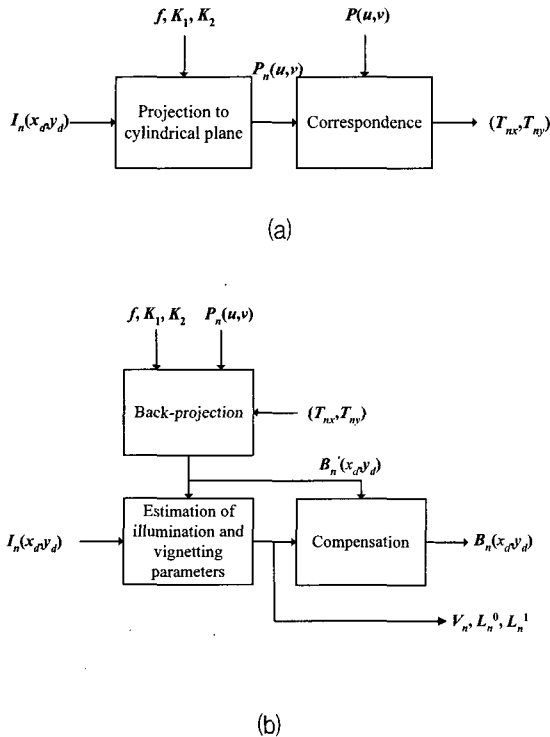


그림 5. (a) 영상 정합 블록도, (b) 파라미터 및 배경 영상 추정 블록도
 Fig. 5. (a) Block diagram for matching and (b) parameter and background estimation.

소화 하는 방법으로 구하게 된다.

$$E_p = \sum_{u_m, v_m}^{\text{overlapped area}} e_p^2(u_m, v_m)$$

$$e_p(u_m, v_m) = \frac{P_m(u_m, v_m)}{V_p \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{u_m^2 + v_m^2}}{f} \right) \right) + 1 - V_p} - \frac{P_{m+1}(u_{m+1}, v_{m+1})}{V_p \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{u_{m+1}^2 + v_{m+1}^2}}{f} \right) \right) + 1 - V_p}$$

$$u_m = u_{m+1} + \Delta u_m, \quad v_m = v_{m+1} + \Delta v_m \tag{5}$$

여기서 $(\Delta u_m, \Delta v_m)$ 는 특징점과 대응점의 차에 대응되는 움직임 벡터를 나타낸다. 연속되는 영상을 효과적으로 합성하기 위하여, 선형 보간을 통하여 파노라믹 영상을 합성하게 된다.

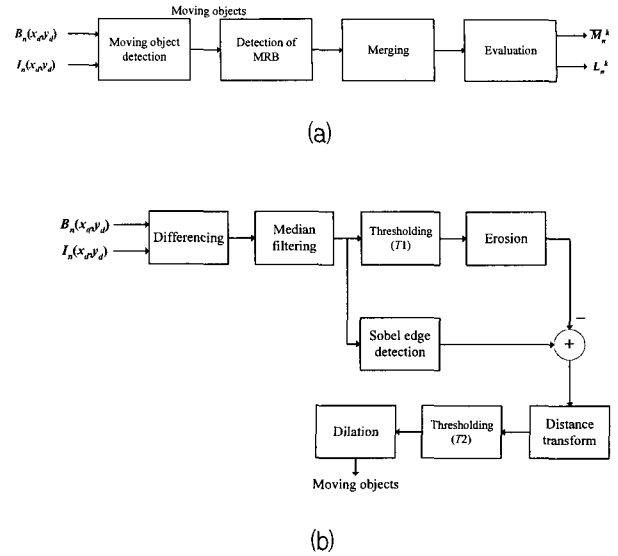


그림 6. (a) 움직이는 물체 지역 검출 블록도, (b) 움직이는 물체 추출 블록도
 Fig. 6. (a) Block diagram for moving rectangular region detection and (b) moving object detection block.

3. 움직임 물체 검출

제안한 파노라믹 영상 모델링 기반 비디오 압축 시스템의 초기화가 끝나고, 실제 동작 모드에서 입력 영상 $I_n(x_d, y_d)$ 는 움직이는 물체, 배경 그리고 차영상으로 분해된다. 배경 영상은 조명보상과 비그네이팅 파라메터 그리고 전체 배경영상인 $P(u, v)$ 에 대응되는 대응점을 사용하여 복원된다. 그림 5(a)는 대응점을 구하기 위한 블록도를 나타낸다. 먼저, 입력 영상이 실린더 평면으로 투영되고, 이것을 $P_n(u, v)$ 로 나타낸다. 이때 투영과정은 전체 파노라믹 배경영상을 예측할 때 사용했던 파라메터를 이용하여 이루어진다. 카메라의 회전은 실린더 영상 평면에서는 평행이동으로 나타나기 때문에, 대응점 (T_{nx}, T_{ny}) 는 다음의 오차값을 최소화함으로써 추정할 수 있다.

$$E = \sum_{\text{all } u, v \in P_n(u, v)} \delta_T(P_n(u, v) - P(u - T_{nx}, v - T_{ny})) \tag{6}$$

여기서 $\delta_T(x)$ 는 x 값이 T 보다 작을 때 0이고, T 와 같거나 클 때 1이 되는 함수를 나타낸다. 여기서 T 값은 식 (5)의 표준편차를 사용하였다. 이 오차 함수는 배경영상으로 고려될 수 있는 화소의 수와 같은 의미로 해석될 수 있다. 본 논문에서는 기존의 많은 예측 방법에서 사용하는 오차의 제곱의 합을 사용하지 않았다. 이것으로 인하여, 움직이는 물체에 의한 outlier를 제거하고 대응



그림 7. 실린더 좌표계 투영 후 파노라믹 배경 영상
 Fig. 7. Panoramic background image obtained after cylindrical mapping.

표 1. 입력 영상 시퀀스
 Table 1. Input video sequences.

Sequence	Size	Acquisition	Contents	Camera parameters	Number of frames
Sequence 1	360x240	Camcorder	Indoor / no moving object	Automatic	157
Sequence 2	360x240	Camcorder	Indoor / no moving object	Fixed	170
Sequence 3	320x240	USB camera	Indoor / no moving object	Automatic	212
Sequence 4	360x240	Camcorder	Outdoor / no moving object	Automatic	169
Sequence 5	360x240	Camcorder	Indoor / moving object	Automatic	1112
Sequence 6	360x240	Camcorder	Indoor / moving object	Fixed	589
Sequence 7	320x240	USB camera	Indoor / moving object	Automatic	671
Sequence 8	360x240	Camcorder	Outdoor / moving object	Automatic	443

점을 강건히 예측할 수 있게 된다^{[17][18]}. 입력 영상과 전체 파노라믹 배경 영상이 크기 때문에, 식 (6)를 최소화 하는 대응점을 찾는 데 많은 계산량을 요구하게 된다. 본 연구에서는 이러한 계산량을 감소시키기 위하여 계측적 방법을 사용하였다.

조명 및 비그네이팅 파라미터와 배경 영상을 복원하기 위하여 그림 5(b)의 블록도를 제안하였다. 여기서 역투영은 $P_n(u,v)$ 을 다시 영상 평면으로 복원 시켜주는 과정을 나타낸다. 역투영은 식 (1)과 (3)에 의하여 생성된 파노라믹 영상에서 사용된 역 변환에 의하여 수행된다. 배경 영상 $B_n(x_d, y_d)$ 는 다음의 식에 의하여 보정이 되지 않은 배경영상에서의 조명 및 비그네이팅 효과를 보정하게 된다.

$$B_n(x_d, y_d) = \frac{B'_n(x_d, y_d) \cdot \left\{ V_n \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{x_d^2 + y_d^2}}{f} \right) \right) + 1 - V_n \right\} - L'_n}{L_n^0} \quad (7)$$

여기서, V_n , L_n^0 과 L_n^1 는 다음의 오차를 최소함으로써 구할 수 있다.

$$E = \sum \rho(I_n - B_n) = \sum \rho \left(I_n - \frac{B'_n(x_d, y_d) \cdot \left\{ V_n \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{x_d^2 + y_d^2}}{f} \right) \right) + 1 - V_n \right\} - L'_n}{L_n^0} \right) \quad (8)$$

여기서, 비용함수는 다음과 같이 정의된다.

$$\rho(x) = \frac{(x/a)^2}{1+(x/a)^2} \quad (9)$$

$$\left(\frac{TB}{f_r} - 64\right) - T_{mo}. \quad (12)$$

이 비용함수는 outlier의 역할을 하는 움직임 물체 영역에 의한 오차를 최소화시킬 수 있다. 여기서 임계치 a는 다음의 식의 표준편차로 설정된다.

$$e_p(u_m, v_m) \times V_p \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{u_{m+1}^2 + v_{m+1}^2}}{f} \right) \right) + 1 - V_p \quad (10)$$

4. 움직임 물체 압축 및 비트율 제어

제안한 파노라믹 영상 모델링 기반 비디오 압축 시스템은 차영상과 움직이는 물체를 JPEG-2000에 의하여 정지영상 압축을 수행하고, 배경영상은 몇 개의 카메라 파라미터와 전체 배경영상을 통하여 합성된다. 각 입력 영상 한 장에 대하여 전체 8바이트의 파라미터를 필요로 한다. 여기서 4 바이트는 대응점 (T_{nx}, T_{ny})를 위한 것이고, 2 바이트는 조명 변화를 위한 파라미터이다. 나머지 2 바이트는 비그네이팅 파라미터를 위한 것이다. 주어진 목표 비트율 TB에 대하여, 각 움직이는 물체에 대하여 다음과 같이 목표 비트가 설정된다.

$$T_{mo} = \left(\frac{TB}{f_r} - 64\right) \times \{TR_{mo} + (1 - TR_{mo}) \times \beta\} \quad (11)$$

여기서, 는 전체 영상에 대하여 움직이는 직사각형 영역들에 의하여 점유되는 지역의 비율을 나타낸다. TR_{mo}는 0에서부터 1의 값을 가질 수 있고, f_r은 프레임 레이트를 나타낸다. 여기서 TR_{mo}는 0.3으로 설정하였다. 차영상을 위하여 설정한 목표 비트는 다음과 같이 정하였다.

III. 실험결과 및 토의

제안한 파노라믹 모델링 기반 비디오 압축 시스템은 기존의 JPEG-2000^[19]과 MPEG-4^[7]와 실제 비디오 시퀀스에 대하여 성능을 평가하였다. 표 1은 본 연구에서 사용한 비디오 시퀀스를 나타낸다. 실험 영상 시퀀스는 소니 비디오 캠코더와 인텔 USB 카메라를 통하여 취득되었다. 다양한 시퀀스의 특징에 대한 실험을 위하여 실내 및 실외 영상 들로 구성되어 있다. 또한, 카메라의 화이트 밸런스와 노출시간 자동 조정을 온/오프 시켜 비디오를 촬영하였다. 실험 영상 중에 두 개의 시퀀스는 자동 모드 조정을 오프로 한 상태에서 취득되었으며, 나머지는 자동 모드 조정 온 상태에서 취득된 영상이다. 자동 모드 조정은 단독 영상 관점에서 인간에게 좋은 화질로 보일 수 있으나, 파노라믹 영상을 합성하기 위하여 주변 영상들과 같이 사용할 때는 칼라의 특성이 달라지는 문제점을 가지고 있다. 표 1에서 본 실험에서 사용한 영상 중에 시퀀스 1과 5는 (2와 6, 3과 7, 4와 8) 같은 장소에서 취득된 것이다. 본 실험에선 제안한 방법과 MPEG-4와 JPEG-2000 알고리즘에 대하여 PSNR과 사용 비트량을 비교하였다. MPEG-4의 경우에는 프레임 레이트를 변경 시키면서 실험을 진행하였다. 실제 보안 카메라 응용에서는 MPEG과 같은 동영상 압축 방식을 사용할 지라도, 낮은 프레임 레이트가 일반적으로 사용되고 있는 실정이다.

표 2. 제안한 방법, JPEG-2000, MPEG-4에 의한 시퀀스 1, 2, 3, 4에 대한 비디오 압축 성능 비교
Table 2. Comparison of the proposed algorithm, JPEG-2000, and MPEG-4 with four video sequences in terms of PSNR and coded data bytes/frame.

	Panoramic modeling only	JPEG-2000	MPEG-4		
			10 frames/s	5 frames/s	1 frame/s
Sequence 1	26.6 (11)	26.2 (587)	29.9 (557)	29.5 (658)	27.9 (1854)
Sequence 2	25.7 (11)	25.3 (455)	29.8 (560)	29.3 (714)	28.2 (1872)
Sequence 3	23.2 (11)	23.6 (389)	29.9 (534)	29.9 (570)	28.5 (1819)
Sequence 4	19.8 (11)	19.6 (501)	26.1 (1598)	26.0 (1882)	25.2 (5899)

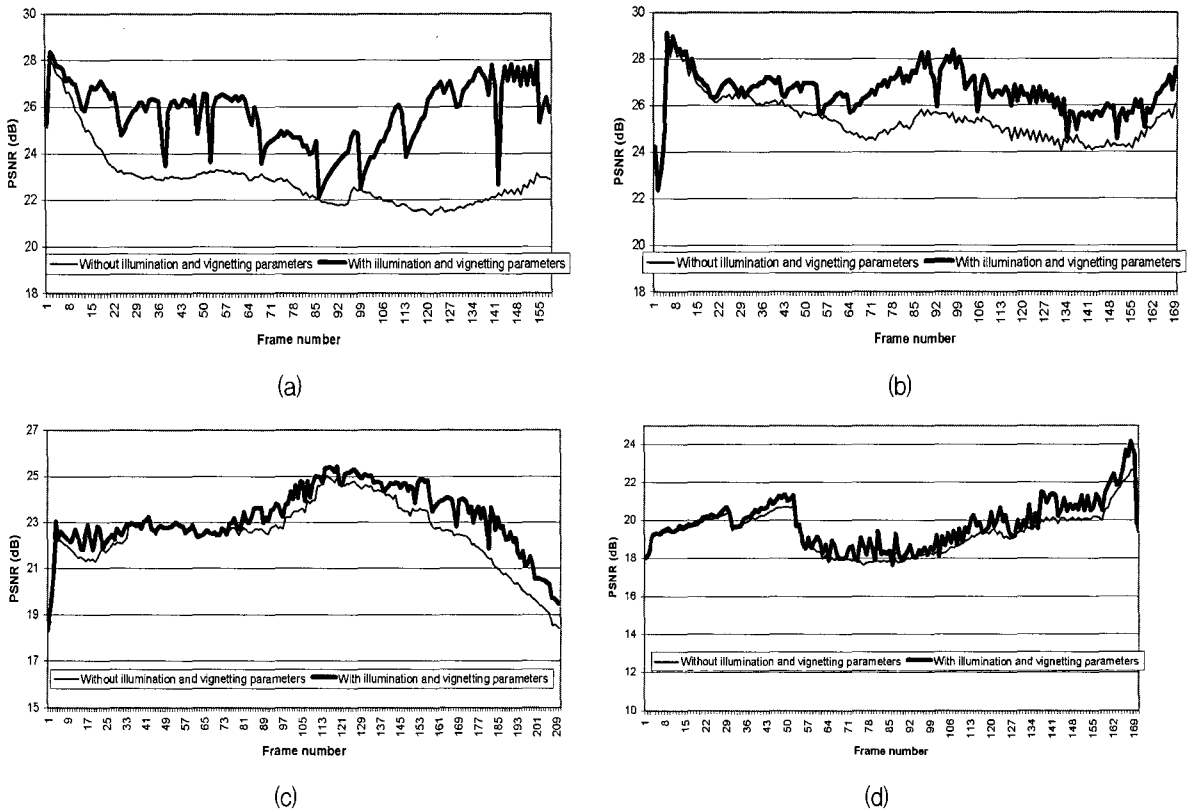


그림 8. 조명 및 비그네이팅 파라미터 보정 및 비보정에 따른 복원 비디오 PSNR (a) 시퀀스 1, (b) 시퀀스 2, (c) 시퀀스 3, (d) 시퀀스 4.

Fig. 8. PSNR of video sequences with and without the illumination and vignetting parameters. (a) Sequence 1, (b) sequence 2, (c) sequence 3 and (d) sequence 4.

표 3. 제안한 방법, JPEG-2000, MPEG-4에 의한 시퀀스 5, 6, 7, 8에 대한 비디오 압축 성능 비교

Tablw 3. Comparison of the proposed algorithm, JPEG-2000, and MPEG-4 with four video sequences in terms of PSNR and coded data bytes/frame

	Panoramic modeling and moving object	Panoramic, moving object, and residual	JPEG-2000	MPEG-4		
				10 frames/s	5 frames/s	1 frame/s
Sequence 5	26.2 (238)	29.4 (580)	26.2 (765)	29.7 (595)	29.4 (718)	28.2 (1873)
Sequence 6	25.1 (265)	28.7 (532)	25.7 (762)	29.5 (606)	29.1 (762)	28.1 (1876)
Sequence 7	23.1 (390)	25.0 (645)	23.7 (679)	29.1 (709)	29.0 (826)	28.2 (1893)
Sequence 8	20.7 (313)	23.1 (612)	19.5 (864)	26.2 (1749)	26.0 (2194)	25.5 (5427)

1. 배경영상 복원

그림 7은 시퀀스 1과 5를 사용하여 복원한 파노라믹 배경영상을 나타낸다. 본 파노라믹 영상 합성은 그림 4의 방법을 통하여 합성된 영상이다. 본 배경 영상은 JPEG-2000에 의하여 압축되고, 복호화 쪽에서 복원될

것이다. 표 2는 움직이는 물체를 포함하지 않고, 배경영상만 반복적으로 나타낼 때, 네 개의 비디오 시퀀스에 대하여 PSNR과 영상당 사용된 바이트 수를 비교하였다. 파노라믹 영상 모델링 기법의 경우 움직이는 물체가 없기 때문에, 단지 11바이트의 정보만으로 배경 영

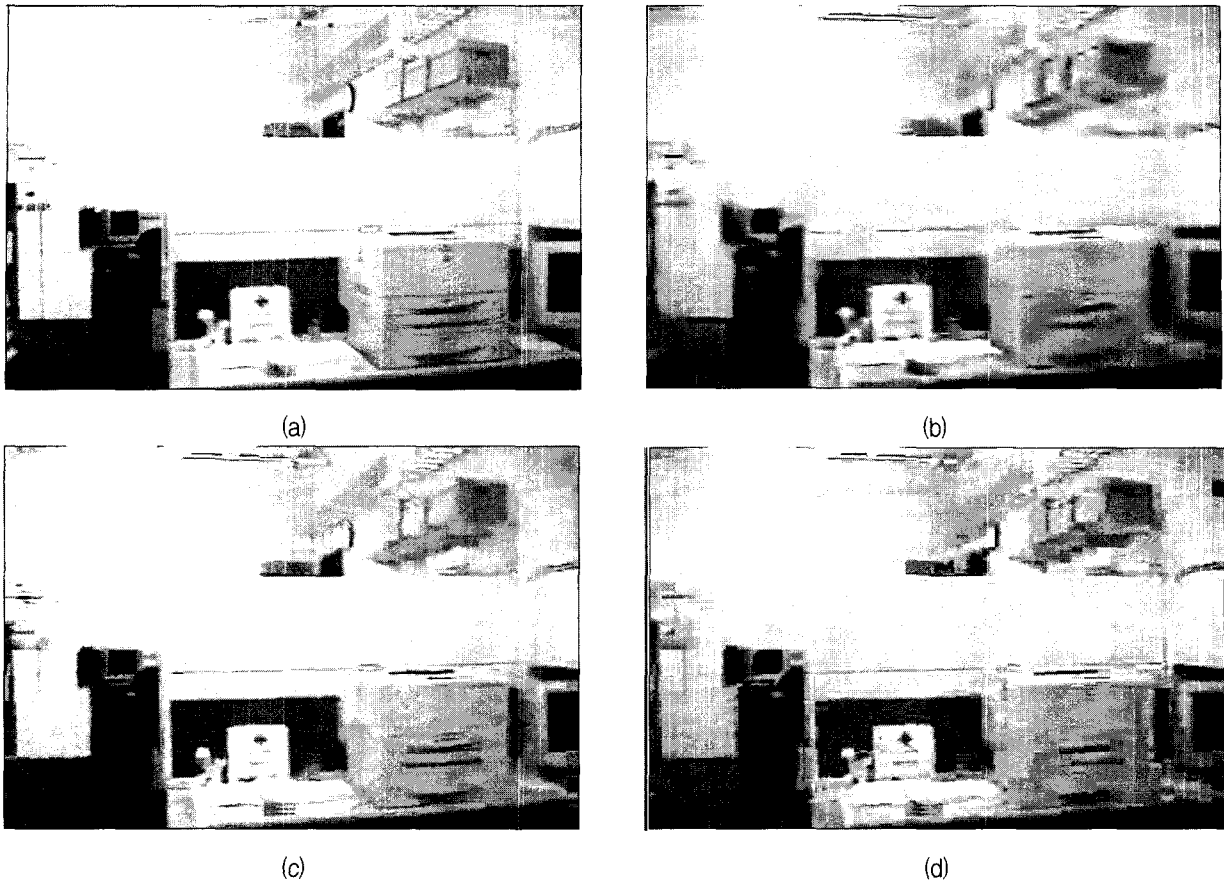


그림 9. 움직이는 물체가 없는 비디오 복원 결과 (a) 제안한 방법 (28.3dB), (b) JPEG-2000 (28.2 dB), (c) MPEG-4, 10 frame/s (29.5 dB) and (d) MPEG-4, 1 frame/s (28.4 dB)

Fig. 9. Reconstructed image without any moving object. (a) Proposed panoramic modeling (28.3 dB), (b) JPEG-2000 (28.2 dB), (c) MPEG-4, 10 frame/s (29.5 dB) and (d) MPEG-4, 1 frame/s (28.4 dB).

상을 복원할 수 있을 것이다. MPEG-4 비디오 압축 방법의 경우는 프레임 레이트가 작아질수록 큰 성능 저하를 볼 수 있다.

그림 8은 조명 보상과 비그네이팅 보정이 있을 경우와 없을 경우에 대한 PSNR을 보여 주고 있다. 제안한 방법에서 이러한 카메라 파라미터를 사용함으로써 1-2 dB 정도의 성능 향상을 관찰할 수 있었다. 그림 9에서 보듯이 같은 PSNR 일지라도 제안한 방법의 주관적 화질은 기존의 방법들에 비하여 상당히 좋음을 확인할 수 있다. 기존의 방법들은 블록 기반이고 예측 변환 기법을 사용함으로써 블록킹 효과 및 ringing 효과 등을 볼 수 있다.

2. 움직이는 물체를 가진 비디오 영상 실험

표 3은 움직이는 물체를 포함하는 네 개의 시퀀스에 대한 PSNR과 프레임당 사용된 바이트를 보여주고 있다. 차영상 압축을 사용할 경우 JPEG-2000 방법에 비하여 거의 같은 PSNR에서 42 ~ 69% 정도의 압축 효율

을 보이고 있다. 제안한 방법의 비디오 압축 효율이 낮은 프레임 레이트에서의 MPEG-4 압축 방법에 비하여 효과적임을 알 수 있다. 또한 차영상을 약 300바이트 정도를 사용하여 압축을 할 경우 PSNR 관점에서 2~3dB 향상됨을 볼 수 있다.

그림 10(a)는 움직이는 물체에 대한 압축 없이 파노라믹 기반 배경 복원과 차영상 압축에 의하여 복원된 영상을 나타낸다. 그림 10(b)는 움직이는 물체 추출 후 압축과 파노라믹 영상 복원을 사용한 경우이다. 그림 10(a)의 PSNR이 그림 10(b)의 것에 비하여 좋은 값을 보여주나, 주관적 화질에서는 그림 10(b)가 명백히 더 우수함을 알 수 있다. 인간이 실제로 움직이는 사물에 더 관심을 가지기 때문이다. 제안한 방법의 경우는 움직이는 물체에만 비트를 할당하는 효과를 가지게 된다. 실제로 움직이는 물체를 포함하는 최소 직사각형 안에서의 제안한 방법의 PSNR은 30.1dB인 반면 그림 10(a)의 것은 25.3dB이다. 그림 10(b)에서 제안한 방법에 의하여 움직임 물체 추출이 모든 경우에 동작되지 않는

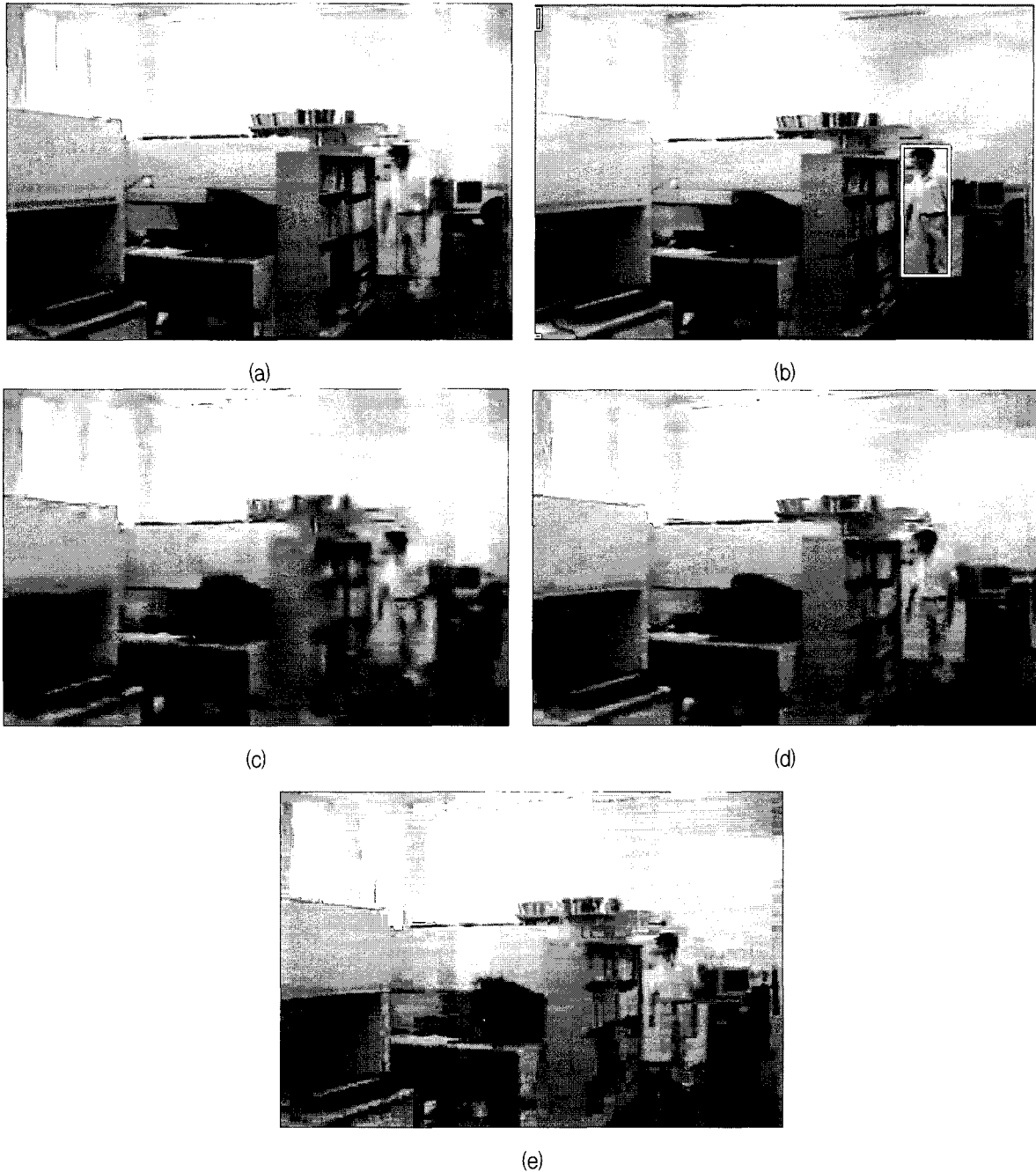


그림 10. 움직이는 물체가 포함된 비디오의 복원 영상 (a) 파노라믹+차영상 (30.2dB, 335), (b) 파노라믹+움직이 물체 (30.2dB, 335), (c) JPEG-2000 (27.2 dB, 977),(d) MPEG-4, 10 frames/s (29.5 dB, 451) and (e) MPEG-4, 1 frame/s (27.8 dB, 1852).

Fig. 10. Reconstructed image with a moving object. (a) Panoramic modeling and residual image coding (30.2 dB, 335), (b) panoramic modeling and moving object coding with no residual coding (27.3 dB, 388), (c) JPEG-2000 (27.2 dB, 977), (d) MPEG-4, 10 frames/s (29.5 dB, 451) and (e) MPEG-4, 1 frame/s (27.8 dB, 1852).

다. 실제로 그림에서 보듯이 움직이는 사람의 다리 부분이 움직이는 물체로 추출되지는 않고 있다. 그러나 이러한 부분은 차영상 압축을 추가함으로써 보상이 되게 된다. 그림 10(c)~(e)는 JPEG-2000과 MPEG-4에 의하여 압축 복원된 영상을 보여주고 있다. 같은 PSNR에

서 주관적 화질에서 많은 차이를 보이고 있으며, 비트율 면에서 제안한 방법이 효과적임을 알 수 있다.

3. Rate-distortion 비교

그림 11은 제안한 방법과, JPEG-2000 그리고

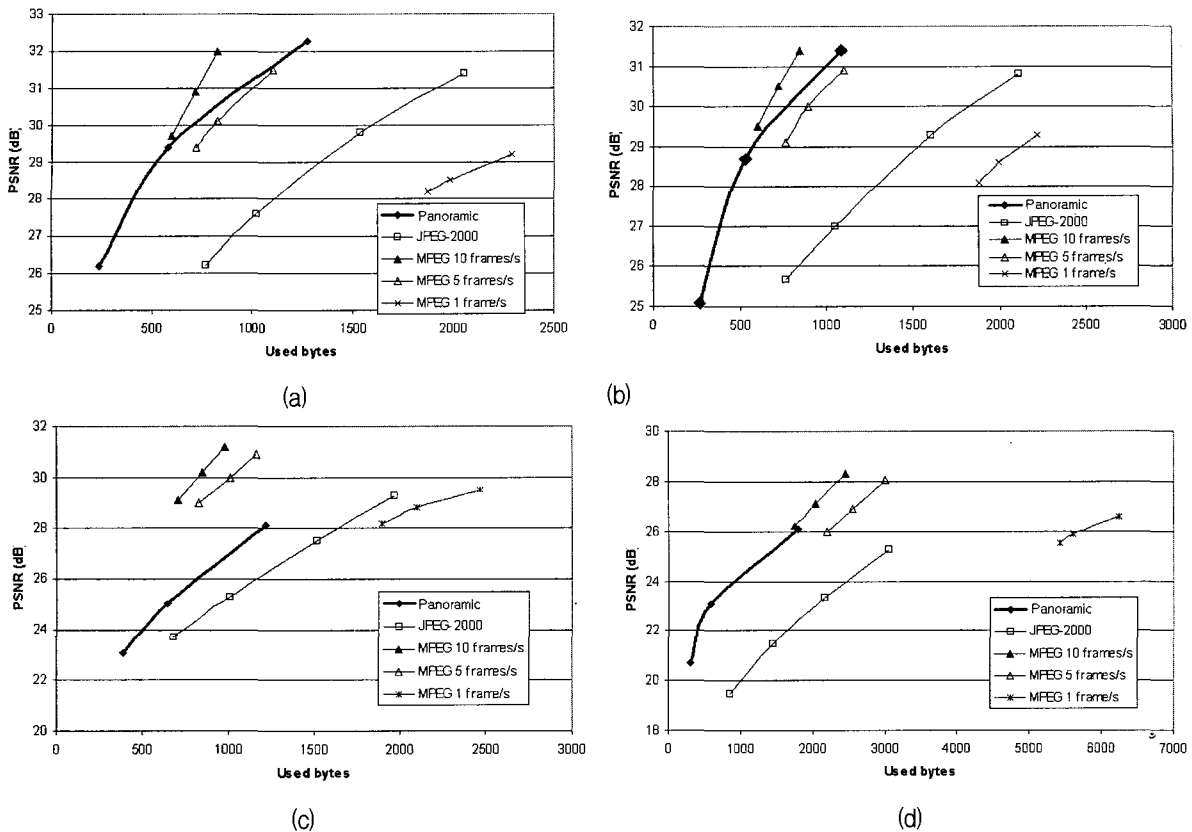


그림 11. 제안한 알고리즘, JPEG-2000과 MPEG-4에 의한 RD 성능 (a) 시퀀스 5, (b) 시퀀스 6, (c) 시퀀스 7, (d) 시퀀스 8.
 Fig. 11. Rate-distortion curves of the proposed algorithm, JPEG-2000, and MPEG-4. (a) Sequence 5, (b) sequence 6, (c) sequence 7 and (d) sequence 8.

MPEG-4에 의한 비트율과 PSNR을 보여 주고 있다. 제안한 방법이 JPEG-2000이나 MPEG-4의 낮은 비트율 압축에서 효과적임을 알 수 있다. MPEG-4 비디오 압축 기법이 높은 프레임 레이트에서 좋은 결과를 보여 주고 있다. 그러나 낮은 프레임 레이트에서는 제일 좋지 않은 결과를 보여 주고 있다. MPEG-4 비디오 압축 기법의 성능은 프레임 레이트에 의하여 크게 영향을 받음을 알 수 있다. 실제 다수의 카메라들을 사용하는 보안 시스템의 경우 낮은 프레임 레이트는 아주 일반적인 상황으로, MPEG과 같은 동영상 압축 방식이 효과적이지 않을 수 있다.

IV. 결 론

본 논문에서는 보안 카메라 응용을 위하여 움직이는 물체 추출과 효과적인 파노라믹 영상 모델링 기반 동영상 압축 방법을 제안하였고, 실제 영상을 사용하여 그 효용성을 보였다. 파노라믹 영상 모델링의 경우, 실제 상황에 근접하는 여러 가지 내부 및 외부 파라미터를

사용하여, 정확한 카메라 모델링을 이루었다. 제안한 방법에 의하여 객관적 성능 향상 뿐만 아니라 주관적 화질 향상이 매우 큼을 확인하였다. 실제로 낮은 프레임 레이트에서의 MPEG-4나 JPEG-2000 압축 방식에 비하여 1~4dB 정도의 성능향상을 얻었다.

참 고 문 헌

- [1] I. Pavlidis, V. Morellas, P. Tsiamyrtzis, and S. Harp, "Urban surveillance systems: From the laboratory to the commercial world," *Proceedings of IEEE*, vol. 89, pp. 1478-1497, 2001.
- [2] C. S. Regazzoni and G. L. Foresti, "Guest editorial: Video processing and communications in real-time surveillance," *Real-Time Imaging*, vol. 7, pp. 381-388, 2001.
- [3] C. Regazzoni, V. Ramesh, and G. L. Foresti, "Special issue on video communications, processing, and understanding for third generation surveillance systems," *Proceedings of IEEE*, vol. 89, pp. 1355-1367, 2001.
- [4] B. Rieger and H. Rode, "Digital image recording

- for court-related purposes," Proc. Security Technology, pp. 262-279, 1999.
- [5] T. Boulton, R. J. Micheals, X. Gao, and M. Eckmann, "Into the woods: Visual surveillance of noncooperative and camouflaged targets in complex outdoor setting," Proceedings of IEEE, vol. 89, pp. 1382-1402, 2001.
- [6] T. Sikora, "The MPEG-4 video standard verification model," IEEE Trans. Circuits and Systems for Video Technology, vol. 7, pp. 19-31, 1997.
- [7] ISO/IEC 14496-2: Information technology-Coding of audio-visual objects, Part-2.
- [8] C. Christopoulos, J. Askelof, and M. Larsson, "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard," Int. Conf. Image Processing, vol. 2, pp. 247-249, 2000.
- [9] ISO/IEC 15444-1: Information technology-JPEG 2000 image coding system-Part 1: Core coding system, 2000.
- [10] H.-Y. Shum and R. Szeliski, "Construction and refinement of panoramic mosaics with global and local alignment," Proc. Conf. Computer Vision, pp. 953-956, 1998.
- [11] R. Szeliski, "Video mosaics for virtual environments," Computer Graphics and Applications, vol. 15, no. 2, pp. 22-30, 1996.
- [12] S. Peleg and J. Herman, "Panoramic mosaics by manifold projection," Proc. Computer Vision Pattern Recognition, pp. 338-343, 1997.
- [13] P. J. Burt, M. Irani, S. C. Hsu, P. Anandan, and M. W. Hansen, "Method and apparatus for performing mosaic based image compression," United States Patent 5,991,444, Nov. 1999.
- [14] Y. Altunbasak, R. M. Mersereau, and A. J. Patti, "A fast parametric motion estimation algorithm with illumination and lens distortion correction," IEEE Trans. Image Processing, vol. 12, pp. 395-408, 2003.
- [15] Z. Zhang, "On the epipolar geometry between two images with lens distortions," Proc. Int. Conf. Pattern Recognition, vol. 1, pp. 407-411, 1996.
- [16] G. Stein, "Lens distortion calibration using point correspondences," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 602-608, 1997.
- [17] M. J. Black and P. Anandan, "The robust estimation of multiple motions: Affine and piecewise smooth flow fields," Computer Vision and Image Understanding, vol. 63, pp. 75-104, 1996.
- [18] D.-G. Sim and R.-H. Park, "Robust reweighted MAP motion estimation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, pp. 353-365, 1998.
- [19] M. D. Adams and F. Kossentini, "Jasper: a software-based JPEG-2000 codec implementation," Int. Conf. Image Processing, vol. 2, pp. 53-56, 2000.

 저 자 소 개



심 동 규(정회원)

1993년 2월 서강대학교 전자공학과 학사 졸업
 1995년 2월 서강대학교 전자공학과 공학 석사
 1999년 2월 서강대학교 전자공학과 공학 박사
 2000년 9월 현대전자 기반기술 연구소
 2002년 3월 바로비전 뉴미디어 연구소
 2005년 2월 University of Washington
 2005년 3월 ~ 현재 광운대학교 컴퓨터공학과

<주관심분야 : 비디오, 신호처리, 컴퓨터비전>