

경계-보존 움직임 추정기법

정회원 최명환*, 임정은*, 손광훈*

Edge preserving motion estimation technique

Myounghwan Choi*, JeongEun Lim*, Kwanghoon Sohn* *Regular Members*

요 약

본 논문에서는 경계영역을 잘 보존하는 움직임 추정기법을 제안하였다. 고정크기 블록으로 움직임 추정을 하는 경우 생길 수 있는 경계영역에서의 왜곡은 인간의 시각에 민감하게 작용할 수 있다. 이는 다른 움직임을 가지는 객체에 대해 한 개의 벡터를 할당하기 때문에 발생한다. 이를 위해 가변크기 블록기반의 움직임 추정기법이 제안되었지만 고정크기 블록기반인 압축표준과 호환되지 않는 문제점을 가지고 있다. 본 논문에서 제안한 움직임 추정기법은 고정크기 블록기반으로서 기본적으로 MAD(Mean Absolute Difference)가 최소가 되도록 하는 동시에 영상의 경계값과 복잡도를 이용하여 경계부분에서 일어나는 시각적인 왜곡을 줄일 수 있도록 하였다. 제안한 움직임 추정기법은 기존의 경계영역 보존 기법에 비해 객관적 성능 및 주관적 화질이 향상됨을 모의 실험결과로부터 확인하였다.

ABSTRACT

In this paper, we propose a block-based motion estimation method considering image edges. The existing block-based motion estimation techniques usually provide poor motion compensation near at edges which are very sensitive to the human visual system. In order to solve the problems near at edges, several variable-size block matching algorithms have been introduced to provide more accurate prediction along moving edges. However, they are not directly compatible with the current standards which employ fixed-size block matching algorithms. The proposed method modifies the cost function of the conventional one to preserve edges well enough by considering the characteristics of the block. Simulation results show that this approach preserves edges quite well with good objective quality.

1. 서론

동영상 압축은 디지털 기술의 중요한 응용분야의 하나로서 HDTV, 주문형 비디오 및 영상 회의 등의 다양한 응용분야에서 제한된 용량의 채널을 통하여 동영상을 전송하거나 저장매체에 저장하기 위하여 사용된다. 움직임 추정 및 보상기법은 동영상에 존재하는 시간적 중복성(temporal redundancy)을 제거하여 높은 압축률을 얻는데 핵심적인 역할을 한다. 움직임 보상 기법은 움직임벡터를 이용하여 보상하는 과정과 원 영상과 움직임이 보상된 영상의 차

신호를 부호화하는 과정으로 이루어진다. 여기서 좋은 화질의 움직임 보상은 신뢰성 있는 벡터를 추정함으로써 얻을 수 있다. 이런 목적으로 동영상 압축 표준인 H.261[1], H.263[2], 및 MPEG[3][4]등에서는 고정 크기 블록 정합 알고리즘(Fixed size Block Matching Algorithm)을 이용하여 움직임 추정 및 보상과정을 수행하고 있다.

FBMA에서는 영상을 일정한 크기의 블록으로 나눈 뒤, 블록 내의 모든 화소들이 같은 방향으로 이동한다는 것을 가정하여 이전 영상의 탐색영역에서 비용함수가 최소가 되는 위치를 찾는다. FBMA는

* 연세대학교 전기·전자공학과 디지털 영상 미디어 연구실(khsohn@yonsei.ac.kr)

논문번호: 010087-0508, 접수일자: 2001년 5월 8일

※ 이 논문은 2001년도 한국 학술진흥재단의 지원에 의하여 연구되었음(KRF-2001-041-E00205)

알고리즘이 간단하여 하드웨어 구현이 용이하고 빠른 수행속도를 갖는다. 하지만 대상 영상의 움직임에 대하여 적절한 탐색영역과 블록의 크기를 결정하기 어렵고, 블록 단위로 움직임을 추정하기 때문에 블록화 현상(blocking artifact)이 발생하며 경계면을 포함한 블록에 대해 움직임 추정하는 경우 물체의 일부가 어긋나 버리는 심각한 시각적인 왜곡을 발생시킬 수 있다.

탐색범위에 대한 움직임 추정의 신뢰도는 다해상도(multiresolution)에 기반한 계층적 구조 움직임 탐색(hierarchical motion estimation)[5]을 이용하여 얻을 수 있고 움직임 보상과정에서 나타나는 블록화 현상의 문제는 중첩된 블록 움직임 보상 방법(Overlapped Block Motion Compensation)[6]을 이용하여 해결할 수 있다. 또한 움직임이 있는 경계영역에 대해 시각적인 왜곡과 블록화 현상을 최소화하기 위해 여러 개의 객체를 가지고 있는 블록을 분할하여 하나의 객체를 하나의 블록으로 표현하는 가변크기 블록 기반의 움직임 추정 기법(Variable size Block Motion estimation Algorithm)[7]이 소개되었다. 그러나 VBMA의 경우 FBMA를 사용하고 있는 표준과 호환되지 않는 문제점을 가지고 있다.

본 논문은 움직임이 있는 경계면에서의 시각적인 왜곡을 줄이는 FBMA기반의 움직임 추정기법을 제안한다. 제안한 움직임 추정기법에서 사용한 비용함수는 영상의 경계값과 밝기값을 동시에 고려하여 평활영역과 텍스처영역에서는 밝기값에 대한 오차가중치를 두어 PSNR을 높게 유지하면서 경계영역에서는 경계값에 대한 오차가 중치를 두어 시각적인 왜곡을 최소화하였다.

본 논문의 구성은 다음과 같다. 제 II 장은 경계영역을 보존하는 기존 기법의 소개 및 문제점 분석을 한다. 제 III 장에서는 제안한 기법을 소개하며, 제 IV 장에서 기존의 기법과 제안한 기법을 실험하여 결과를 분석하고 끝으로 제 V 장에서 결론을 맺고자 한다.

II. 기존의 경계영역 보존기법

블록기반 움직임 추정은 구현이 용이하고 효율성면에서 우수하기 때문에 H.261, MPEG과 같은 압축표준에 사용되었다. 그러나 한 블록안에 있는 모든 화소값은 움직임이 같다는 가정을 하여 여러 객체에 대해 한 개의 벡터로만 나타내기 때문에 여

러 가지 문제점이 생길 수 있다. 그림 1은 움직이고 있는 탁구공과 탁구채를 움직임 보정한 영상으로 블록정합 후 생기는 경계영역에서의 시각적인 왜곡을 보여주고 있다. 그림 1(b)의 경우 공의 경계부분이 잘못 보상되어졌는데 이것은 잡음성의 비균일 화소값을 가지고 있는 배경이 블록의 대부분을 차지하여 움직임 벡터를 결정하고 상대적으로 적은 영역을 차지하고 있는 탁구공의 왼쪽부분에 같은 움직임 벡터가 적용이 되었기 때문에 나타난다.

경계영역에서 정확한 벡터를 찾기 위한 기존의 방법은 영상을 고정크기의 블록으로 나눈 후 움직임이 있고 경계영역의 블록이라고 판단될 때 식(1)과 같이 경계값에 대한 오차인 MADE(Mean Absolute Difference of Edge value)를 사용하여 움직임 추정을 함으로써 경계면에 대한 주관적 화질을 향상시키고 그 밖의 다른 영역에서는 식(2)와 같은 밝기값에 대한 오차인 MADI(Mean Absolute Difference of Intensity value)를 사용하여 움직임 추정을 수행한다^[8].

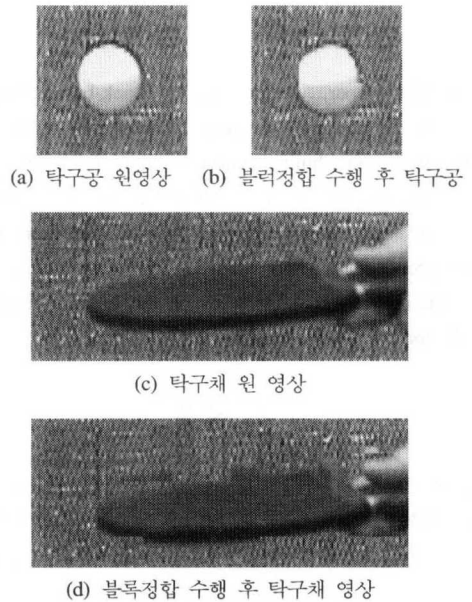


그림 1. 블록정합 수행 후 생기는 경계영역문제 ('TABLE TENNIS' 영상)

$$MADE = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |S_n(i, j) - S_{n-1}(i+x, j+y)| \quad (1)$$

$$MADI = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |I_n(i, j) - I_{n-1}(i+x, j+y)| \quad (2)$$

여기서 $MADI$ 는 현재 영상 I_n 과 이전 영상 I_{n-1} 에 대한 밝기값의 차이를 말하고 $MADE$ 는 Sobel 경계검출기로 얻은 경계영상 S_n 과 S_{n-1} 에 있는 경계값의 차이를 말한다. x, y 는 찾고자하는 움직임 벡터를, N 은 블록의 크기를 나타낸다

현재 영상은 각 블록마다 그림2와 같은 과정으로 진행되는데 움직임 검출은 식(3)을 이용하여 구하게 된다.

$$Fd_n = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |I_n(i, j) - I_{n-1}(i, j)| \quad (3)$$

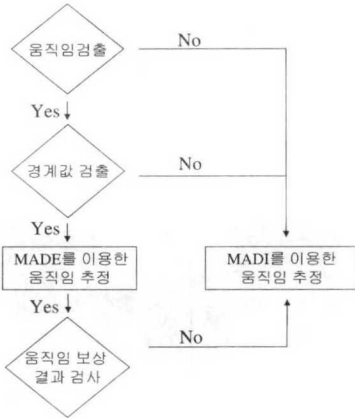
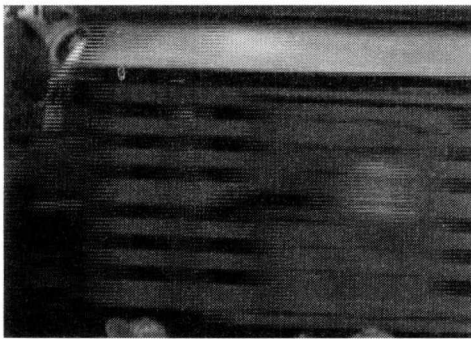
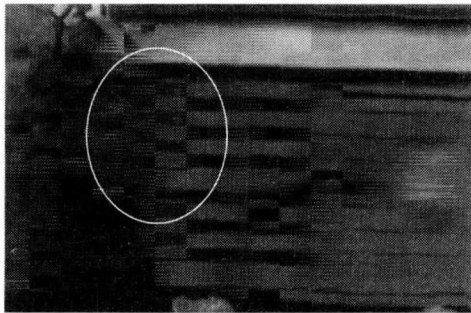


그림 2. 기존 기법의 흐름도



(a) 원영상



(b) 기존기법을 이용해 보정한 영상

그림 3. MADE를 이용한 추정의 문제점

여기서 $I_n(i, j)$ 은 현재 프레임의 밝기값, $I_{n-1}(i, j)$ 은 이전 프레임의 밝기값, N 은 블록의 크기를 말한다. 식(3)에서 구한 Fd_n 이 Th_{fd} 보다 크면 움직임이 있는 블록으로 판단하고 작으면 움직임이 없는 블록으로 설정한다. 다음 과정으로 움직임이 있는 블록이 경계영역을 포함하는지에 대해 경계값을 이용해 검사하게 된다. 경계영역 판단은 현재 블록안에 있는 경계값의 개수로서 판단을 하게 되는데 이때 잡음으로 인한 작은 경계값들은 Th_e 를 이용하여 무시한다. 경계영역을 포함한 블록으로 판단되면 $MADE$ 를 이용해 움직임 추정을 수행 함으로써 경계영역을 잘 보존할 수 있다. 또 움직임 보상이 잘 되었는지 판단하기 위해 $MADE$ 값이 Th_{MADE} 보다 클 경우 $MADE$ 를 이용한 움직임 추정이 실패한 것으로 판단하고 $MADI$ 를 이용해 다시 추정한다. 그러나 $MADE$ 로 움직임 추정을 하는 경우 그 위치에서의 밝기값에 대한 오차의 고려없이 단지 경계영상에서 경계값에 대한 오차만을 고려하기 때문에 복잡한 영역이나 텍스처 영역에 대해 찾은 벡터의 신뢰성이 떨어진다. 즉 $MADE$ 값이 작더라도 그 위치에 대한 $MADI$ 는 클 수 있어서 텍스처영역과 같이 일정밝기 값이 반복되어 배치되어 있는 영역인 경우 그림3(b)와 같은 심각한 오류를 발생시킬 수 있다.

이를 위해 비용함수 $MADE$ 를 이용해 찾은 벡터가 신뢰성이 있는지를 판단할 때 $MADE$ 대신 $MADI$ 값을 사용 함으로써 텍스처영역에서 발생할 수 있는 시각적인 왜곡을 막을 수 있을 것이다. 본 논문에서는 영역판단에 따라 $MADE$ 또는 $MADI$ 를 독립적으로 사용하기보다는 적절한 가중치와 함께 두 비용함수의 합으로 이루어지는 새로운 비용함수를 사용한 움직임 추정기법을 제안한다.

III. 경제-보존 움직임 추정기법

본 논문에서 제안한 움직임 추정기법은 블록의 경계값과 복잡도를 이용하여 비용함수를 조절함으로써 경계영역에 대한 시각적인 왜곡을 줄일 수 있도록 하였다. 비용함수는 $MADI$ 와 $MADE$ 로 구성이 되며 블록의 영역에 따라 가중치를 조절한다. 즉 평활영역과 텍스처영역에서는 $MADI$ 에 가중치를 두어 움직임 추정에 신뢰성을 두며 경계영역에서는 $MADE$ 에 가중치를 두어 경계면에 대한 시각적인 왜곡을 줄이도록 하였다. 경계값은 원영상에 저역통

과 필터를 적용한 후 식(4)의 Sobel 경계검출기를 이용하여 구한다. 이로써 MADE는 고주파를 가지는 잡음에 덜 민감하게 되어 움직임 추정에 대한 신뢰성을 향상시킬 수 있게 된다.

$$g_x = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, g_y = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad (4)$$

본 논문에서 제안한 움직임 추정기법에 사용하는 비용함수는 식(5) 및 식(6)과 같다.

$$Cost = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |I_n(i, j) - I_{n-1}(i+x, j+y)|}{(M_{edge} + 1)} + \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |S_n(i, j) - S_{n-1}(i+x, j+y)|}{Func(NZC)} \quad (5)$$

$$Func(NZC) = |NZC - a| + |NZC - b| + c \quad (6)$$

식(5)에서 $I_n(i, j), I_{n-1}(i, j)$ 의 MAD는 MADI를, $S_n(i, j), S_{n-1}(i, j)$ 의 MAD는 MADE를, M_{edge} 는 현재 블록안에 있는 경계값들의 평균[9]을, 그리고 식(6)의 $Func(NZC)$ 는 NZC(number of zero crossing)[10]의 함수로서 블록의 복잡도를 표현한다. 우선 NZC를 구하기 위해 블록안에 있는 화소값들에 대해 그것들의 평균을 빼고 난 후 블록의 행, 열 그리고 두 개의 대각방향을 따라 이웃하는 화소에 대해 부호의 변화를 살핀 후 부호변화가 일어난 횟수를 세어 구한다. 또한 NZC가 잡음에 민감하지 않도록 임계값 Th_{NZC} 를 두어 이 값보다 작은 값은 잡음으로 판단하고 '0'으로 지정함으로써 이웃하는 화소끼리의 부호변화가 일어나지 않도록 하여 불필요한 NZC값의 증가를 방지하였다. 따라서 경계영상에 대해 NZC를 구하면 그림 4와 같이 값의 크기에 따라 현재 블록이 평활 영역, 경계 영역, 텍스처 영역중 어느 영역인지 알 수 있을 것이고 영역에 따라 $Func(NZC)$ 로서 식(5)의 비용함수에 대해 가중치 조절을 하였다.

식(5)는 기본적으로 경계영역에서 M_{edge} 가 증가하므로 MADE의 가중치가 증가하여 시각적인 왜곡을 줄일 수 있는 추정을 하게 된다. 여기서 MADE의 독립적인 사용보다 MADI와 함께 사용함으로써 보다 정확한 움직임 추정을 할 수 있다. 그러나 경계영역이 아닌 경우는 MADI에 가중치를 두어 신뢰성 있는 벡터를 찾을 수 있다. 여기서 경계값으로 가중치를 조절할 경우 경계영역과 평활영역에서는

좋은 결과를 기대할 수 있지만 텍스처영역일 경우도 M_{edge} 가 증가하여 MADE의 가중치가 증가한다. 즉, 제 II 장에서 언급한 기존의 방식과 같은 문제점이 생겨 화질의 열화를 초래할 것이다. 따라서 $Func(NZC)$ 를 이용하여 텍스처영역에 대한 화질열화를 줄일 수 있도록 하였다. $Func(NZC)$ 는 그림 4와 같은 함수로서 텍스처영역에서 값이 증가하여 MADI에 대한 가중치를 증가시켜 잘못된 추정을 막을 수 있도록 하였다. 마찬가지로 평활영역에서는 MADI에 가중치를, 경계영역에서는 MADE에 가중치를 두도록 하였다. 또한 식(6)의 '+1'은 경계값이 없을 때 비용함수가 무한대로 발산하는 것을 막기 위해 필요하다.

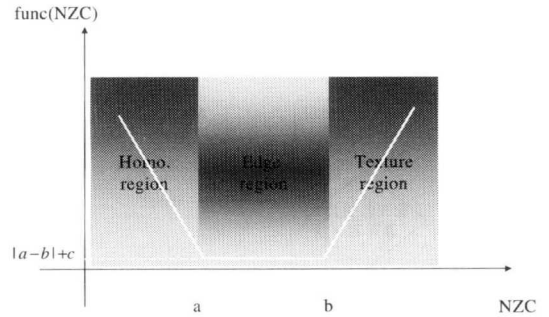


그림 4. $func(NZC)$

IV. 모의 실험 결과 및 분석

본 논문에서 제안한 움직임 추정기법의 성능을 평가하기 위해 모의실험을 수행하였다.

본 실험에서는 'TABLE TENNIS'(720×480: #7 ~ #22), 'TRAIN AND TUNNEL'(720×576: #7 ~ #22), 'CLAIR'(352×288: #2 ~ #20)을 실험영상으로 사용하였다. 움직임 추정을 위한 블록의 크기는 16×16으로 하였고, 탐색영역은 -16~16으로 설정하였으며 $Func(NZC)$ 에 사용한 파라미터들은 실험적으로 결정하여 $a=70, b=110, c=-39$ 그리고 $Th_{NZC} = 3$ 을 사용하였다. 그림 5의 'TABLE TENNIS' 결과 영상과 그림 6의 'CLAIR' 결과 영상을 이용하여 주관적 화질을 비교하였다. 그림 7, 그림 8은 그림 5 'TABLE TENNIS'영상의 확대 영상으로서 손톱, 손목부분 그리고 팔 부분에 대해 제안한 기법이 일반 블록 정합과 기존 기법인 Edge matching에 비해 경계면 부근에서의 주관적 화질이 향상되었음을 확인할 수 있다. 이는 제안 기법이 경계 부분에

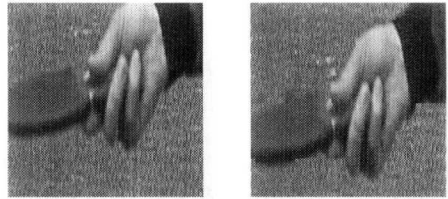
서 경계값 뿐만 아니라 밝기값을 함께 고려함으로써 보다 정확한 벡터를 찾을 수 있기 때문이다. 또한 그림 6 'CLAIR' 영상의 확대 영상인 그림 9의 눈 부분, 그림 10의 입 부분 영상에서 제안 기법의 주관적 화질 향상을 다시 한번 확인할 수 있다.

표 1은 여러 영상 시퀀스에 대해 일반 블록정합, 기존 기법(Edge matching), 제안 기법으로 움직임 보상한 영상에 대한 평균 PSNR을 나타낸다. 기존 기법의 경우 경계 부근에서는 경계값만으로 추정하고 텍스처 영역에서도 신뢰도 있는 추정이 어렵기 때문에 PSNR의 저하가 발생한다. 'TABLE TENNIS' 영상에서 제안 기법이 기존기법에 비해 0.2 dB 정도의 성능향상을 보였으며 'TRAIN AND TUNNEL' 영상에서도 약 0.8 dB의 성능향상을 확인하였다. 또한 그 밖의 영상에서도 기존 기법에 비해 제안한 움직임 추정기법이 객관적 화질 면에서 우수한 화질 보임을 확인할 수 있었다. 따라서 제안 방식은 기존의 경계보존 기법보다 영상의 객관적 화질이 향상되었으며 동시에 일반 고정블록 기반 움직임 추정기법과 기존기법에서 생기는 경계 면에서의 시간적인 왜곡을 최소화하였음을 확인하였다.

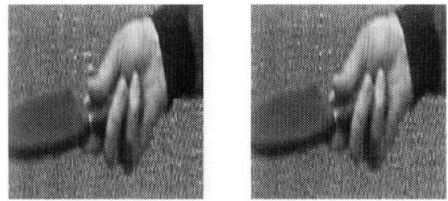


(c) 기존 기법 (d) 제안 기법

그림 6. 'CLAIR'#11 결과영상

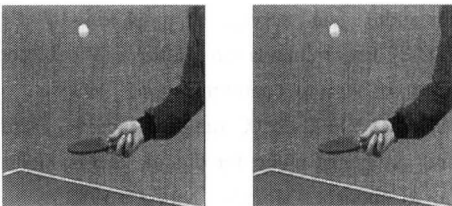


(a) 원 영상 (b) 일반 블록정합

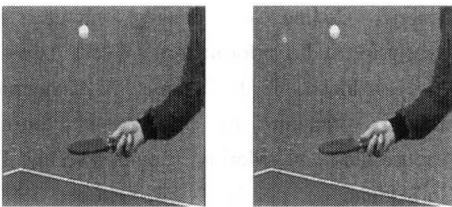


(c) 기존 기법 (d) 제안 기법

그림 7. 'TABLE TENNIS'#5 탁구채 부분 결과영상

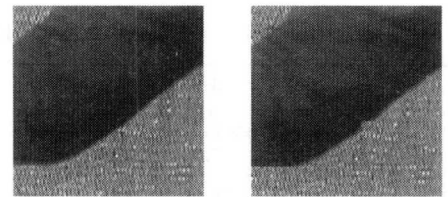


(a) 원 영상 (b) 일반 블록정합

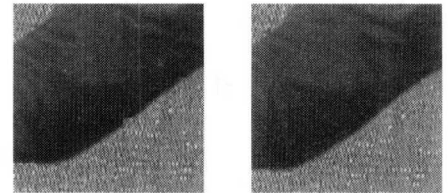


(c) 기존기법 (d) 제안 기법

그림 5. 'TABLE TENNIS'#5의 결과영상 비교



(a) 원 영상 (b) 일반 블록정합

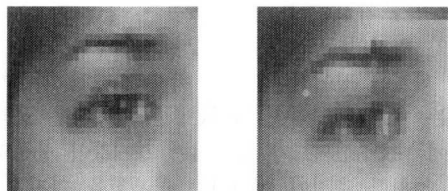


(c) 기존기법 (d) 제안 기법

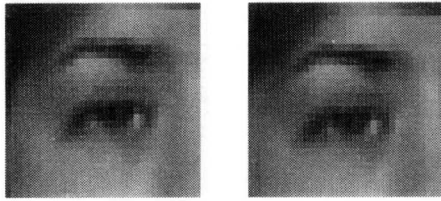
그림 8. 'TABLE TENNIS'#5 팔 부분 결과영상



(a) 원 영상 (b) 일반 블록정합

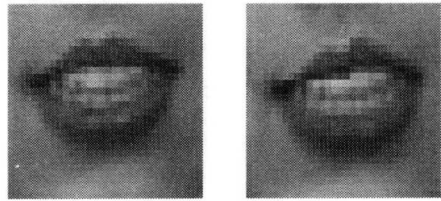


(a) 원 영상 (b) 일반 블록정합

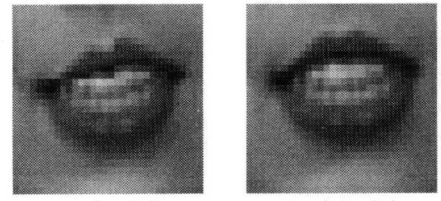


(c) 기존기법 (d) 제안 기법

그림 9. 'CLAIR'#11 눈 부분 결과영상



(a) 원 영상 (b) 일반 블록정합



(c) 기존기법 (d) 제안 기법

그림 10. 'CLAIR'#11 입 부분 결과영상

표 1. 각 알고리즘별 움직임 보상 결과 비교(PSNR, dB)

	블록 정합	기존 기법	제안 기법
TABLE TENNIS	27.50	26.84	27.05
TRAIN AND TUNNEL	28.45	27.09	27.90
CLAIR	39.61	39.47	39.50

V. 결론

본 논문에서는 인간 시각에 민감한 경계영역을 효과적으로 보존하기 위한 움직임 추정기법을 제안하였다.

밝기값만을 고려하는 일반 블록정합의 경우 밝기 값 오차가 가장 작은 곳을 찾는 추정을 하므로 객관적인 화질인 PSNR은 높지만 경계부근에서의 주관적 화질이 떨어진다. 이를 위해 소개된 기존기법에서는 텍스처 영역을 가지는 블록의 경우 움직임 추정에 대한 신뢰도가 떨어지며 경계값이 불분명한 경우 정확한 추정이 어렵다는 문제점을 가지고 있다.

제안 기법은 압축표준과 같은 구조인 고정크기 블록기반으로써 영상에 있는 여러 정보를 이용하여 보다 정확한 추정이 가능하도록 하였다. 기본적으로 영상의 밝기값과 경계값을 동시에 이용하는 움직임 추정을 하며 현재 블록의 경계값 평균과 복잡도를 이용하여 가중치 조절을 함으로써 경계영역인 경우 밝기값은 고려하지만 경계값을 중심으로 하는 추정이 되도록 하고 평활 영역과 텍스처 영역에서는 밝기값을 중심으로 하는 추정이 되도록 하였다. 제안한 기법은 기존의 경계보존 움직임 추정기법보다 우수한 PSNR을 얻을 수 있었고, 주관적으로도 다른 기법들 보다 경계부근에서의 시각적인 왜곡을 최소화하였음을 확인할 수 있었다.

결론적으로 제안기법은 경계값의 평균을 이용하여 경계부근에서의 시각적인 왜곡을 최소화하고 평활 영역과 텍스처 영역에서는 블록의 복잡도를 이용하여 기존기법에서 발생할 수 있는 오류를 줄이도록 하였다.

참고 문헌

- [1] ITU-T Recommendation H.261, "Video codec for audio visual services at p×64 kbits/s," 1990.
- [2] ITU-T Recommendation H.263, "Video codec for low bit rate communication," 1995.
- [3] ISO/IEC 11172-2, "Coding of moving Pictures and associated audio for digital storage media at up to about 1.5Mbits/s : video," 1992.
- [4] ISO/IEC 13818-2, "Information technology - Generic Coding of moving picture and associated Audio information : video," 1994.
- [5] M. H. Ahmad, T. J. Dennis, "A hierarchical motion estimator for interframe coding," Applications of Motion Compensation, IEE Colloquium on, pp. 3/1 - 3/6, October, 1990.
- [6] S. Nogaki, M. ohta, "An overlapped block motion compensation for high quality motion picture coding," Circuits and Systems, 1992. ISCAS '92. Proceedings., 1992 IEEE International Symposium on, vol. 1, pp. 184 -187, May, 1992.
- [7] M. H. Chan, Y. 'B. Yu, and A. G. Constautimides, "Variable size block matching motion compensation with applications to video coding," IEE Proceedings. Communications,

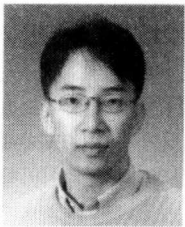
Speech and Vision, vol. 137, no. 4, pp. 205-212, August, 1990.

[8] Y. L. Chan, and W. C. Siu, "Edge oriented block motion estimation for video coding," IEE Proceedings. Visual Image Signal Process, vol. 144, no. 3, pp. 136-144, June, 1997.

[9] M. H. Kim and K. H. Sohn, "Edge preserving directional regularization technique for disparity estimation of stereoscopic image," IEEE Transactions on Consumer Electronics, vol. 45, no. 3, pp. 804-811, August, 1999.

[10] Kundu, "System and method for matching blocks in a sequence of images," US PATENT, no. 5974192, October, 1999.

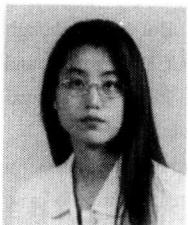
최 명 환(Myounghwan Choi)



-2000년 : 한양대학교 전자컴퓨터 공학과, 공학사
 -2002년 : 연세대학교 전기전자공학과 석사
 -현재 : LG 전자 DTV 연구소 연구원

<주관심 분야> 3차원 영상 부호화/복호화, 영상 압축, DTV

임 정 은(JeongEun Lim)



-1999년 : 경원대학교 전자공학과, 공학사
 -2001년 : 연세대학교 전기전자공학과 석사
 -현재 : 연세대학교 전기전자공학과 박사과정

<주관심 분야> 3차원 영상 부호화/복호화, 영상 압축, 영상 통신

손 광 훈(Kwanghoon Sohn)

한국 통신학회 논문지 제 25권 7A호 참조