

# 적응적 변이추정 기법을 이용한 새로운 중간시점영상합성

정희원 배 경 훈\*, 김 은 수\*

## A New Intermediate View Reconstruction using Adaptive Disparity Estimation Scheme

Kyung-hoon Bae\*, Eun-soo Kim\* *Regular Members*

### 요 약

본 논문에서는 적응적 변이추정 기법을 이용한 새로운 중간시점영상합성 방법을 제시하였다. 즉, 스테레오 입력 영상으로부터 특징값을 추출하고, 설정된 임계값과 비교하여 특징값의 크기를 결정한 다음, 특징값의 크기에 따라 정합창의 크기를 적응적으로 선택하여 정합함으로써 중간시점의 영상을 합성하는 새로운 기법을 제안하였다. 제안된 기법에서는 배경과 같이 특징값이 작은 영역에서는 비교적 큰 정합 창에 의한 블러정합이 이루어지고 물체의 윤곽선과 같이 특징값이 큰 영역에서는 상대적으로 작은 정합창에 의한 미세한 정합이 적응적으로 수행되기 때문에 전체적인 정합성능의 개선뿐만 아니라 기존의 기법에서 나타나는 오정합이나 블록화 현상등의 문제점 해결의 가능성을 제시하였다. 또한, 'Parts' 및 'Piano' 영상을 사용한 실험결과 본 논문에서 새로이 제안한 중간시점 영상 합성 방법은 기존의 방식에 비해 평균적으로 PSNR은 약 2.32~4.16dB가 향상되었고, 수행시간은 약 39.34~65.58% 감소됨을 확인하였다.

### ABSTRACT

In this paper, a new intermediate view reconstruction technique by using a disparity estimation method based-on the adaptive matching window size is proposed. In the proposed method, once the feature values are extracted from the input stereo image, then the matching window size for the intermediate view reconstruction is adaptively selected in accordance with the magnitude of this feature values. That is, coarse matching is performed in the region having smaller feature values while accurate matching is carried out in the region having larger feature values by comparing with the predetermined threshold value. Accordingly, this new approach is not only able to reduce the mismatching probability of the disparity vector mostly happened in the accurate disparity estimation with a small matching window size, but is also able to reduce the blocking effect occurred in the disparity estimation with a large matching window size. Some experimental results on the 'Parts' and 'Piano' images show that the proposed method improves the PSNR about 2.32~4.16dB and reduces the execution time to about 39.34~65.58% than those of the conventional matching methods.

### I. 서론

일반적으로, 사람이 물체를 입체적으로 볼 수 있는 것은 양안시차에 의한 것으로, 두 눈은 각기 다

른 시점의 영상을 관측하고 인간의 뇌가 이 두 스테레오 영상의 차이를 합성하여 3차원(3D) 입체물체를 인식하게 된다<sup>[1]</sup>. 그 동안, 이러한 인간시각 시스템(HVS: human visual system)을 모방하여 다양

\* 광운대학교 전자공학과 국가지정 3차원 미디어 연구실(khbae@explore.kwangwoon.ac.kr)

논문번호 : 020084-0222, 접수일자 : 2002년 2월 22일

※ 본 연구는 과학기술부 국가지정 연구실(NRL) 사업의 지원 아래 수행되었습니다.

한 형태의 양안식 3D 디스플레이 시스템이 구현되어 왔다. 그러나 이러한 스테레오 양안방식은 시점이 2안으로 제한되기 때문에 관찰자가 제한된 시역을 벗어나거나 초점이 맞지 않을 경우에는 입체감을 느낄 수 없고, 또한 눈의 피곤함 및 어지러움 등을 느낄 수 있기 때문에 실질적인 응용이 제한되고 있다<sup>[2]</sup>. 따라서 이러한 기존의 스테레오 방식의 단점을 해결하기 위한 접근방법으로 여러 형태의 다시점 스테레오 3D 디스플레이 시스템에 관한 연구들이 활발히 진행되고 있다<sup>[3]</sup>. 이러한 시스템은 다안식 입체 카메라를 통해 다시점의 영상획득과 디스플레이가 이루어지기 때문에 시점의 개수가 증가됨에 따라 시역이 확대되고 보다 자연스러운 3D 디스플레이가 가능하다. 그러나 시점의 개수가 증가됨에 따라 영상데이터도 동시에 증가하기 때문에 실시간의 영상처리 프로세서 및 초고속, 광대역의 전송채널이 요구되게 된다<sup>[4]</sup>. 최근, 이러한 문제점을 해결하기 위한 새로운 접근방법으로 제한된 개수의 스테레오 영상만을 이용하여 여러 개의 다시점 스테레오 영상을 디지털적으로 합성해내는 중간시점영상합성(intermediate view reconstruction)기법이 연구개발 되고 있다<sup>[5][6]</sup>. 이러한 중간시점 영상합성 기법은 제한된 스테레오 영상으로부터 입체의 다시점 스테레오 영상을 디지털적으로 합성해 내기 때문에 기존의 3D 디스플레이 시스템의 문제점 해결과 동시에 시역이 확대된 자연스러운 3D 입체 디스플레이의 구현 가능성을 제시하고 있다.

일반적으로 중간영상합성 과정에서 변이정합 방법으로는 화소기반 정합 알고리즘(pixel-based matching algorithm), 블록기반 정합 알고리즘(block-based matching algorithm) 및 특징기반 정합 알고리즘(feature-based matching algorithm)등의 방식과 메쉬(mesh)를 이용하는 정합 방식 등이 사용되는데 이 중에서, 화소기반의 정합방식은 세밀한 정합(dense matching)은 가능하지만 정합과정에서 오정합(mismatching)이 나타날 확률이 높고 수행시간이 길다는 등의 단점이 있다. 또한, 블록기반 정합방식의 경우는 수행시간은 짧으나 세밀한 정합이 어렵고 블록화 현상(blocking effect)등의 단점이 나타나고 있다<sup>[7]</sup>. 그리고 특징기반 정합방식은 입력영상의 특징에 따른 효율적인 정합이 가능하지만 특징값이 없는 유사영역 부분에서 잘못된 변이할당이 생기는 등의 단점이 있다. 또한 메쉬 기반 방식은 블록화 현상이 없이 효율적인 부호화가 가능하지만 잘못된 절점의 추출로 잘못된 변이 추정이 발

생할 수 있다.

따라서 본 논문에서는 이러한 단점들이 보완된 적응적 정합창에 기반한 새로운 변이추정 기법을 이용한 중간시점영상합성 방법을 제안하고자 한다. 즉, 제안된 방법에서는 스테레오 입력영상으로부터 특징값을 추출하고, 설정된 임계값(threshold value)과 비교하여 특징값의 크기를 결정한 다음, 그 특징값의 크기에 따라 정합창의 크기를 적응적으로 선택하여 정합을 수행하게 된다. 결과적으로, 제안된 기법은 스테레오 입력영상에서 배경과 같이 특징값이 작은 영역에서는 비교적 큰 정합창에 의한 변이정합을 수행하고, 물체영상의 윤곽선과 같이 특징값이 큰 영역에서는 상대적으로 작은 정합창에 의한 미세한 정합을 적응적으로 수행하게 된다. 이를 통해 전체적인 합성영상의 성능개선 뿐만 아니라 기존의 기법에서 나타나는 오정합, 블록화 현상 및 과도한 수행시간 등의 문제점 해결의 가능성을 제시하고자 한다.

그리고 실험을 통해 본 논문에서 새로이 제시된 적응적 변이추정에 기반한 중간시점 영상합성 기법의 성능을 기존 방법과 비교 분석하고자 한다. 즉, 서로 특징이 다른 'Parts' 및 'Piano' 영상을 실험 영상으로 사용하여 제안한 방법과 기존의 방법을 각각 이용하여 중간시점 영상을 합성한 다음, 그 결과들을 PSNR과 수행시간의 관점에서 서로 비교 분석하고자 한다. 또한, 실험 결과를 통해 본 논문에서 새로이 제안한 기법의 우수성을 제시함으로써 중간시점 영상을 이용한 실질적인 다시점 3D 디스플레이 시스템의 구현 가능성을 제시하고자 한다.

## II. 적응적 변이추정 방식

그림 1은 본 논문에서 새로이 제시한 적응적 변이추정에 기반한 중간시점영상합성의 전체 흐름도를 나타낸 것이다. 본 논문에서는 특징기반의 변이추정을 이용하여 스테레오 입력영상으로부터 특징값을 추출하고<sup>[8][9]</sup> 설정된 임계값과 비교하여 특징값의 크기를 결정한 다음, 그 특징값의 크기에 따라 적응적으로 정합창의 크기를 선택하여 정합을 수행하게 된다. 또한, 보다 자연스러운 정합을 위하여 가중 평균값에 의한 내삽(interpolation with weighted mean value) 방법으로 중간시점 영상을 합성하였다<sup>[6]</sup>. 그리고 중복 할당된 변이나 할당되지 않은 변이에 대해서는 변이 안정화 과정을 거쳐 그 주변영역 변이의 평균값으로 대체하게 된다. 이상의 중간시점

영상합성 과정을 통해 보다 자연스러운 3D 입체감을 갖는 다시점의 중간시점 영상을 합성하게 된다.

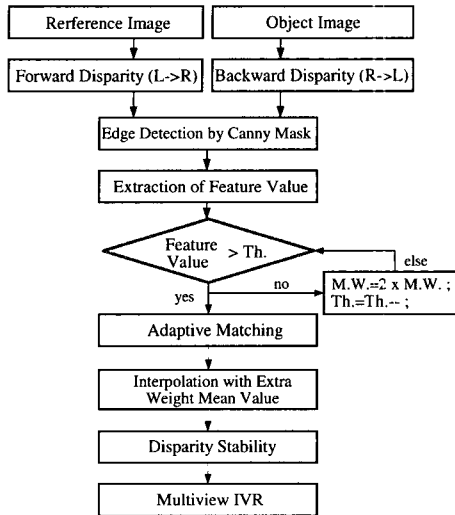


그림 1. 제안된 중간영상합성 알고리즘의 전체흐름도

1. 스테레오 입력 영상의 특징값 추출

일반적으로, 스테레오 입력영상으로부터 특징값을 추출할 때 사용되는 단방향 정합방식(one-directional matching)은 한 방향에서 기준영상과 가장 유사한 영역만을 찾아내어 정확한 정합점을 찾기가 어렵기 때문에 본 논문에서는 단방향이 아닌 양방향 정합방식 (bi-directional matching)을 통해 특징값을 추출하였다<sup>10)</sup>. 또한, 특징기반의 변이추정 방식을 이용하여 특징값을 추출할 때 그림 2와 같이 비교적 잡음에 덜 민감한 윤곽선 검출 미분연산자인 캐니 마스크 (Canny mask)를 사용하였다.

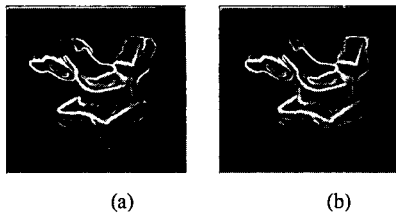


그림 2. 캐니 마스크 연산자를 적용한 'Part'의 (a)좌영상 (b)우영상

그리고 스테레오 입력영상으로부터 변이를 추출하기 위해 사용된 비용함수로는 계산량은 좀 늘어나지만 보다 정확한 변이 추출을 위해 MSE(mean square error)를 사용하였다. 식 (1)은 스테레오 영상의 변이를 추출하기 위해 사용된 MSE 함수를 나

타낸 것이며, 식에서  $N_x, N_y$ 는 블록 또는 화소의 크기를,  $I_L$ 과  $I_R$ 은 좌측영상 블록과 이에 대응하는 우측영상 블록을 각각 나타낸 것이다.

$$\frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |I_L(i, j) - I_R(i + k, j)|^2 \quad (1)$$

그림 3의 '복도 영상'은 에피플라 선상에서 좌, 우 영상간에 서로 대응되는 블록을 찾는 과정을 예로 나타낸 것이다. 즉, 좌측 영상을 일정한 크기의 블록으로 나눈 후 임의의 블록에 대해 이에 대응하는 같은 에피플라 선상에 존재하는 우측영상에서의 일치점을 일정한 범위 ( $\pm S_x$ )내에서 찾는 과정을 보여준 것이다. 일반적으로, 인간시각 시스템에서는 수직변이가 수평변이에 비해 그 영향이 적기 때문에<sup>2)</sup> 이러한 인간 시각 시스템의 특성을 고려하여 병렬로 스테레오 카메라를 구성하게 되면 수평변이만이 존재하기 때문에 그림 3과 같이 수평 성분으로 변이 벡터를 찾을 수 있다.

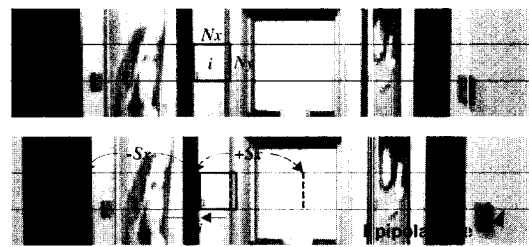


그림 3. 에피플라 선 상의 대응 블록 탐색도

식 (2)는 그림 2에서 표현된  $\pm S_x$ 의 범위를 나타낸 것이다. 여기서  $S$ 는 우측영상에서 대응되는 블록을 찾는 탐색범위를,  $x_0, y_0$ 는 좌측영상의 블록의 시작점을 각각 나타낸다.

$$S = \{R(x, y) | (x_0 - S_x \leq x \leq x_0 + S_x + N_x), (y_0 \leq y \leq y_0 + N_y)\} \quad (2)$$

따라서 탐색범위  $\pm S$  내에서 MSE를 최소로 하는 블록을 좌측영상과 정합된 우측영상의 블록으로 선택하게 되면, 우측영상과 기준(좌측) 영상간의 좌표값 차이로 그 블록을 대표하는 변이값을 얻게 된다. 즉, 좌측영상을 기준으로 한 우측영상과의 변이는 식 (3)과 같이 주어지게 된다.

$$\arg \min_{d_x} \frac{1}{N_x N_y} \sum_{j, n} |I_R(m, n) - I_L(m + d_x, n)| \quad (3)$$

여기서,  $I_R$ 는 예측 영상을,  $I_L$ 는 기준 영상을,  $d_{ix}$ 는 탐색범위 내의 변이 값을 각각 나타내며,  $\beta_i$ 는  $I_R$ 의  $i$ 번째  $N_x \times N_y$  크기의 블록을 나타낸다. 식 (3)에서 얻어지는 변이 벡터를 블록을 대표하는 값으로 변이 지도(disparity map)를 나타내는데 사용된다.

## 2. 특징값 크기에 따른 적응적 정합창 선택 알고리즘

그림 4는 스테레오 입력 영상으로부터 추출된 특징값의 크기에 따라 정합될 정합창의 크기를 적응적으로 선택하는 알고리즘의 흐름도를 나타낸 것이다. 먼저, 스테레오 입력영상으로부터 추출된 특징값들을 정규화 하게 된다. 그리고 0부터 1 사이 값으로 정규화 된 특징값의 크기를 결정하기 위해 각 단계별로 임계값이 설정되게 된다. 일단, 임계단계 및 각 단계별 임계값이 설정되면 각 임계 단계에 대응하는 같은 수의 정합창이 설정되게 된다. 본 논문에서는 임계값을 5단계 0.1, 0.3, 0.5, 0.7, 0.9로 설정하였고 이에 대응하는 5가지 크기의 정합창은  $16 \times 16$ ,  $8 \times 8$ ,  $4 \times 4$ ,  $2 \times 2$ ,  $1 \times 1$ 로 설정하였다. 여기서,  $32 \times 32$  크기 이상의 정합창에서는 정합효율이 크게 떨어지기 때문에 5단계의 정합창 중에서 가장 큰 정합창의 크기를  $16 \times 16$ 으로 제한하여 설정하였다. 이상과 같이 임계값과 정합창이 설정되면 그림 4와 같은 과정을 통해 임의의 입력영상으로부터 추출된 특징값에 따라 정합창의 크기가 적응적으로 선택되게 된다. 즉, 정합창 선택 알고리즘의 초기조건에서 임계값은 임계값 중 가장 큰 값인 5단계(0.9)로 설정하여 특징값들과 비교되게 된다. 따라서 비교를 통해 이 임계값 보다 큰 특징값들은 갖는 영역들은 설정된 정합창 중에서 가장 작은 정합창인  $1 \times 1$ 이 선택되어 가장 세밀한 정합이 수행된다. 이 때, 5단계의 임계값 보다 작은 특징값들에 대해서는 임계값을 1단계 감소시킨 4단계(0.7)로 설정한 후 다시 비교된다. 여기서, 임계값보다 큰 특징값을 갖는 영역들에 대해서는 다음 단계의 정합창인  $2 \times 2$ 로 설정되어 정합이 이루어지게 된다. 또한, 이 임계값 보다 작은 특징값들은 앞에서와 같은 방법으로 임계값을 다시 1단계 감소시킨 3단계(0.5)의 임계값과 다시 비교되게 된다. 이때, 임계값 보다 큰 특징값을 갖는 영역들은 다음 단계의 정합창인  $4 \times 4$ 가 설정되어 정합이 이루어지게 된다. 이러한 과정은 임계값이 5단계에서 1단계까지 감소와

증가를 반복하면서 정합창이 적응적으로 설정되게 된다. 따라서 본 논문에서 새로이 제안한 적응적 정합창 선택 알고리즘은 스테레오 입력 영상에서 추출된 특징값의 크기를 설정된 임계값과 비교, 판단하여 적응적으로 적절한 정합창을 설정하게 된다. 즉, 배경과 같이 특징값이 작은 영역은 유사영역으로 판별하여 큰 단위의 정합창을 선택하여 정합을 하고 물체의 윤곽선 같이 특징값이 큰 부분은 화소 단위의 미세 정합을 수행하게 된다. 이를 통해 작은 블록 단위의 미세 정합에서 자주 발생하는 오정합을 감소시킬 뿐만 아니라 큰 블록단위로 정합할 때 발생하는 블록화 현상도 감소시킬 수 있다.

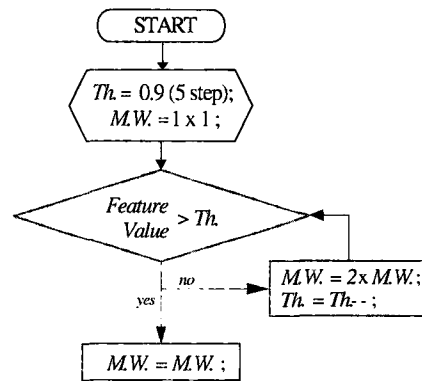


그림 4. 특징값에 따른 정합창 선택 알고리즘 흐름도

## 3. 내삽 기법을 이용한 중간시점 영상합성

그림 5는 일반적인 중간시점 영상합성의 개념도를 나타낸 것이다<sup>[6]</sup>. 스테레오 입력영상으로부터 중간영상을 합성할 때 시점의 위치는  $\alpha$ 로 나타내며,  $\alpha$ 는 0과 1 사이의 값으로 설정되게 된다. 여기서, 좌영상을 기준으로 할 때,  $\alpha=0$ 인 영상은 좌영상을,  $\alpha=1$ 인 영상은 우영상을 각각 나타내게 된다. 중간시점의 영상을 합성하는 방법에는 크게 외삽법(extrapolation)과 내삽법(interpolation)이 있는데 이중, 가중 평균값으로 내삽하는 기법이 외삽법의 경우보다 자연스러운 중간시점영상을 합성할 수 있기 때문에 본 논문에서는 중간시점에 해당하는 영상값은 가중평균값으로 내삽하였다<sup>[5][6]</sup>.

식 (4)는 시점의 위치  $\alpha$ 에 따라 가중 평균값으로 내삽하는 경우를 나타낸 식이다. 여기서,  $d_{ix}$ 는 탐색범위 내의 변이값을,  $I_L$ 는 좌, 우 영상의 변이가 교차되는 지점에 합성된 중간시점 영상의 위치를 각각 나타낸 것이다.

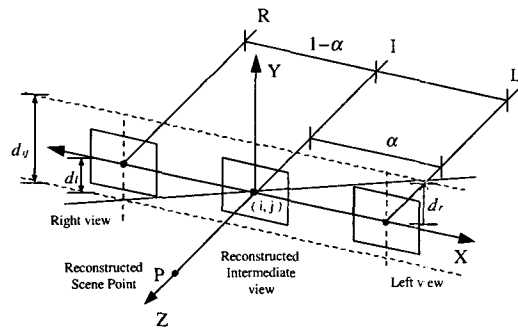


그림 5. 좌, 우 영상 면을 기준으로 한 대응 중간시점

$$L_r(i, j) = (1-\alpha)L_r(i - \hat{d}_y(i, j), j) + \alpha L_l(i - \hat{d}_y(i, j), j) \quad (4)$$

일반적으로, 중간시점 영상을 합성할 때 스테레오 영상쌍의 한쪽 영상에서는 서로 가려진 영역 (occluded region)이 존재하게 된다<sup>[5]</sup>. 이 때, 이러한 가려진 영역에는 변이가 할당되지 못한 영역이 존재할 수 있기 때문에 이러한 영역에 대해서는 변이 안정화 과정을 통하여 주변 변이 값들의 평균값을 구하여 그 변이 값으로 대체하게 된다.

따라서 본 논문에서는 스테레오 입력영상의 특성에 따라 적응적으로 정합창을 선택하여 중간영상을 합성하는 새로운 방식을 제안함으로써 기존 방식들의 장점들을 상호보완적으로 사용하여 성능을 개선하였으며, 이를 이용하여 다시점 3D 디스플레이 시스템을 구현하게 되면 보다 자연스러운 3D 영상 재현의 가능성을 제시하였다.

### III. 실험 및 결과고찰

본 실험에서 사용한 테스트 영상은 그림 6 과 같이 512×512 크기의 'Parts' 영상과 720×576 크기의 'Piano' 영상을 사용하였고<sup>[11][12]</sup>, 컴퓨터 모의실험은 'Microsoft Visual Studio 6.0'에서 수행하였다. 그리고 중간영상 합성방법으로는 기존의 화소기반 정합방법과 화소기반 및 특징기반 방식을 함께 고려한 정합방법과 본 논문에서 새로이 제안한 방식을 실험을 통해 성능을 비교 분석하였다. 성능 비교는 주어진 원 중간영상과 각각의 방법으로 합성된 중간시점영상에 대하여 PSNR과 수행시간 관점에서 비교하였다. 실험에서 입력 영상에 대한 특징값들은 정규화 되고, 임계치는 0.1, 0.3, 0.5, 0.7, 0.9인 5 단계로 설정하였으며, 이를 실험에서는 정수단위의 1부터 5까지의 단계값으로 적용하였다. 초기 조건에

서 임계치는 5단계(0.9)로 설정하였고 그 조건을 만족하지 않으면 임계치가 1 단계만큼 감소되도록 하였다. 초기 임계치는 가변적이고 임의의 설정이 가능하다. 본 논문에서는 초기 임계치를 두 영상 모두 5 값으로 설정하였다. 또한, 이에 대응하는 정합창의 크기는 16×16, 8×8, 4×4, 2×2, 1×1등 5단계로 설정하였다.

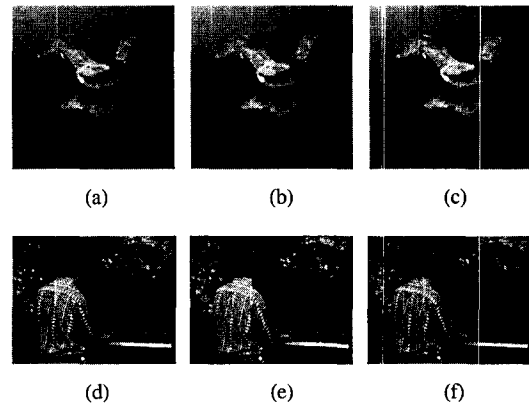
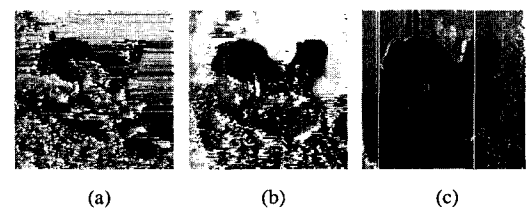


그림 6. 'Parts' 및 'Piano' 원 영상, (a)-(c)'Parts'영상, (d)-(f)'Piano'영상

그림 7은 각 정합방식에 의한 수행되어 얻은 변이 벡터 결과를 변이 지도로 나타낸 것이다. 먼저, 'Parts' 영상의 변이 지도에서 보면, 그림 7(a)의 화소기반 방식은 정밀한 정합을 수행하였으나 물체의 윤곽선 같은 부분에서는 정확한 정합이 이루어지지 않고 있음을 볼 수가 있고, 그림 7(b)의 화소기반 및 특징기반을 같이 고려한 방식에서는 그림 7(a)의 화소기반 방식보다 물체의 윤곽선을 비교적 정확히 추출하고 있음을 볼 수 있지만 그림 7(c)의 제안된 방식과 비교하여 배경 같은 부분에서 제안된 방식이 좀 더 안정된 변이 벡터가 나타남을 볼 수 있다. 또한, 'Piano' 영상에서는 그림 7(d)의 화소기반 방식에서 잘못된 변이 할당이 많이 일어나고 있음을 볼 수가 있고, 그림 7(e)의 화소기반 및 특징기반 방식을 함께 고려한 방식에서는 그림 7(f)과 비교하여 그 차이가 크지는 않지만 그림 7(f)의 제안된 방



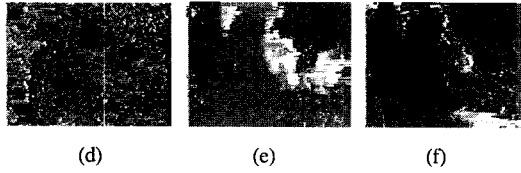


그림 7. 각 기법에 의한 변이지도 (a),(d) 화소기반, (b),(e) 화소 및 특징기반, (c),(f) 제안된 기법

식의 변이 지도가 가장 적은 변이를 갖고 있어 이러한 안정된 변이 벡터를 통해 성능이 보다 개선된 중간영상의 합성결과를 예측 할 수 있다.

그림 8은 각 정합방식에 의한 수행된 중간시점영상 합성결과를 나타낸 것이다. 그림 8에서 보면, 화소기반의 변이추정 방식은 정밀한 정합으로 영상의 세밀한 표현은 가능했지만 변화가 적은 배경 같은 부분에서도 과도한 정밀 정합으로 정합 효율이 떨어지고, 물체의 윤곽선 부분에서는 정확한 윤곽선을 검출하지 못하여 그림 8의 (a)와 (d)와 같이 물체의 윤곽선 부분에서 오정합이 나타남을 볼 수 있다. 그림 8의 (b)와 (e)의 화소 및 특징기반 방식에서는 화소기반 방식에 비해 비교적 물체의 윤곽선 부분의 합성이 잘 수행되었으나 그림 8(c)와 (f)의 제안된 방식 결과와 비교했을 때 겉으로는 큰 차이가 없어 보이지만, 그림 8(d), (e), (f)의 원 표시가 된 귀 근처의 확대영상을 보면 제안된 방식의 합성영상이 기존 방식보다 정확한 정합이 수행되었음을 볼 수 있다.

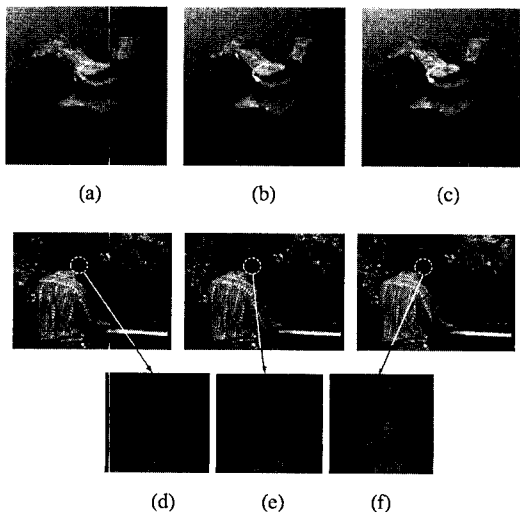


그림 8. 각 기법에 의한 중간영상합성결과, (a),(d) 화소기반, (b),(e) 화소 및 특징기반, (c),(f) 제안된 기법

그림 9는 각 기법으로 합성된 중간영상과 원 영상의 오차영상을 나타낸 것이다. 그림에서 보면 'Parts' 및 'Piano' 영상 모두 본 논문에서 새로이 제안한 방식이 가장 원 영상과 차이가 적음을 알 수가 있다. 이는 제안된 기법으로 합성된 중간영상의 경우 원 영상과 거의 차이가 나지 않다는 것을 의미하며 결과적으로 제안된 방법이 기존 방법 보다 효율적인 중간영상합성기법으로서 그 응용 가능성을 제시한 것이다. 특히, 제안된 방식의 결과 중 그림 9(c)의 경우가 그림 9(f) 보다도 기존의 방식과의 비교해서 현저한 성능차이가 나타남을 볼 수 있다. 이는 제안된 방식이 복잡한 영상으로 구성된 'Piano' 영상보다 비교적 단순한 배경을 가진 영상인 'Parts' 영상이 보다 잘 적용됨을 알 수 있다.

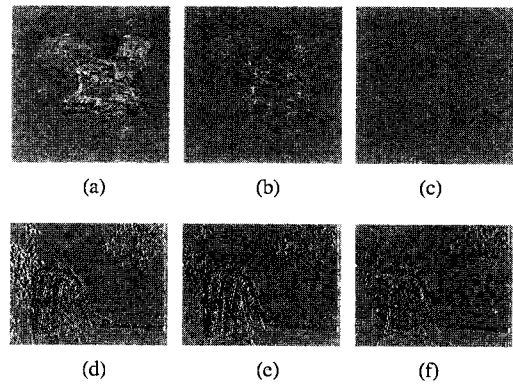


그림 9. 각 기법에 의한 오차영상, (a),(d) 화소기반, (b),(e) 화소 및 특징기반, (c),(f) 제안된 기법

표 1은 원 영상을 기준으로 제안된 방식과 기존의 방식으로 각각 합성된 중간시점 영상의 PSNR을 구한 결과를 나타낸 것이다. 표에서 보면, 제안된 방식이 화소기반의 정합 방식에 비해 'Parts'의 경우는 4.88dB가 'Piano'의 경우는 3.26dB가 각각 개선됨을 볼 수 있다. 또한, 특징 및 화소기반 방식에 대해서는 'Parts'는 2.50dB, 'Piano'는 2.18dB 가 각각 개선됨을 볼 수 있다. 즉, 이상의 PSNR 실험 결과를 보면, 제안된 방식이 기존의 방식 보다 합성 영상의 성능이 크게 개선됨을 알 수 있다.

표 1. PSNR 결과 비교

| Matching Method              | Test Image | PSNR     |          |
|------------------------------|------------|----------|----------|
|                              |            | "parts"  | "piano"  |
| Pixel-based Method           |            | 24.10 dB | 21.82 dB |
| Pixel & Feature-based Method |            | 26.48 dB | 22.90 dB |
| Proposed-Method              |            | 28.98 dB | 25.08 dB |

표 2는 각 기법을 사용하여 중간영상을 합성하는데 걸리는 수행시간을 제안된 방식과 기존의 방식을 비교한 결과를 나타낸 것이다. 먼저, 제안된 방식의 수행시간은 화소기반 방식을 기준으로(100%), 화소기반 방식에 비하여 'Parts'의 경우는 79.43%가, 'Piano'에 대해서는 51.73%가 각각 단축되었다. 또한, 화소 및 특징 기반 방식을 기준으로(100%), 화소 및 특징기반 방식과 비교해서는 'Parts'의 경우가 72.51%, 'Piano'의 경우가 33.61%로 각각 단축됨을 볼 수 있었다. 이상의 실험을 통해 'Parts' 및 'Piano'영상에 대해 제안된 방식이 기존 방식보다 수행시간이 훨씬 짧아짐을 알 수 있으며 특히, 배경과 같이 특징값이 작은 부분을 많이 분포한 'Parts' 영상에서는 상대적으로 큰 정합창이 선택되어 정합이 수행되기 때문에 특징값이 큰 부분이 많이 포함된 복잡한 'Piano' 영상에 비해 수행속도가 훨씬 개선됨을 할 수 있다.

표 2. 수행시간의 결과 비교

| Matching Method              | Test Image | Process Time |         |         |         |
|------------------------------|------------|--------------|---------|---------|---------|
|                              |            | "parts"      |         | "piano" |         |
| Pixel-based Method           |            | 100 %        | 133.65% | 100 %   | 137.54% |
| Pixel & Feature-based Method |            | 74.81%       | 100 %   | 72.71%  | 100%    |
| Proposed-Method              |            | 20.57%       | 27.49%  | 48.27%  | 66.39%  |

#### IV. 결론

본 논문에서는 적응적 변이추정 기법을 이용한 새로운 중간시점 영상합성 방법을 제안하였다. 즉, 스테레오 입력영상으로부터 특징값을 추출하고 설정된 임계값과 비교하여 특징값의 크기를 결정한 다음, 특징값의 크기에 따라 정합 창 크기를 적응적으로 선택하여 정합함으로써 중간시점의 영상을 합성하는 새로운 기법을 제안하였다. 제안된 기법에서는 배경과 같이 특징값이 작은 영역에서는 비교적 큰 정합창에 의한 블록정합이 이루어지고 물체의 윤곽선과 같이 특징값이 큰 영역에서는 상대적으로 작은 정합창에 의한 미세한 정합이 선택적으로 수행되기 때문에 전체적인 정합성능의 개선뿐만 아니라 기존의 기법에서 나타나는 오정합이나 블록화현상 등의 문제점 해결이 가능하다. 또한, 'Parts' 및 'Piano' 두 영상을 사용한 실험결과 본 논문에서 새로이 제안한 중간시점 영상합성 방법은 기존의 방식에 비해 평균적으로 PSNR이 약 2.32~4.12dB 향상되고, 수행시간도 약 39.34~65.58% 감소됨을 확

인하였다. 따라서 본 논문에서 제안한 적응적 정합창에 기반한 새로운 중간영상 합성기법은 앞으로 보다 자연스러운 3D 디스플레이 시스템의 구현에 응용이 가능하리라 판단된다.

#### 참고 문헌

- [1] Victor S. Grinberg, Gregg Podner, M. W. Siegel, "Geometry of binocular imaging", *Proc. of SPIE*, vol. 2177, pp. 56 - 65, February 1994.
- [2] 김은수/이승현 공역, 3차원 영상의 기초, 技多社, 1998.
- [3] N. A. Dodgson, J. R. Moore and S. R. Lang, "Multi-view autostereoscopic 3D display", *IBC '99 (International Broadcasting Convention)*, 10th-14th, pp. 497-502, September 1999
- [4] Jens-Rainer Ohm and Karsten Muller, "Incomplete 3D representation of video objects for multiview applications", *IEEE Trans. on Circuits and Systems for Video Tech.*, vol.9, no.2, March 1999
- [5] Jeffrey S. McVeigh, M. W. Siegel, and Angel G. Jordan "Intermediate view synthesis considering occluded and ambiguously referenced image regions", *Signal Processing: Image Communication*, vol. 9, pp. 21-28, February 1996
- [6] Anthony Mancini, "Disparity estimation and intermediate view reconstruction for noble applications in stereoscopic video", Master thesis, *McGill University*, February 1998
- [7] Andre Redert, Emile Hendriks and Jan Biemond, "Correspondence Estimation in Image Pairs" *IEEE Signal Processing Magazine*, special issue on 3D and stereoscopic visual communication, vol. 16, no.3, pp.29-46, May 1999
- [8] W. E. L. Grimson, "Computational experiments with a feature based stereo algorithm," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.7, no.1, pp.17-34, January 1985.
- [9] W. Ho and N. Ahuja, "Surfaces from stereo: Integrating feature matching, disparity

