

論文2002-39SP-2-12

양방향 시차 몰핑을 이용한 중간 시점 영상 보간

Intermediate Scene Interpolation using Bidirectional Disparity

金大顯*, 尹溶仁*, 崔宗秀*, 金帝佑**, 崔炳胡**

(Dae-Hyun Kim, Yong-In Yoon, Jong-Soo Choi, Je-Woo Kim, and Byeong-Ho Choi)

요약

본 논문은 한 쌍의 스테레오 영상으로부터 BDM(Bidirectional Disparity Morphing)을 이용하여 중간 시점 영상을 보간하는 방법을 제안한다. 영상은 하나 이상의 계층으로 구성되어 있고 각각의 계층은 유사한 시차값을 갖기 때문에, 블록 단위의 시차 맵을 이용하는 것은 유용하다. 또한 블록 단위의 시차 맵은 시차 계산을 위한 영상의 해상도를 줄여 계산량을 크게 감소시킨다. 그러나 서로 다른 계층이 겹쳐진 영역에서는 시차가 급격히 변화하여 정합 오차가 증가하고 가려짐이 발생할 확률이 높다. 따라서 정합 결과에 따라 블록의 크기를 변화시켜 보다 정밀하게 대응 블록을 검색하여 잘못된 시차 계산을 막고, 아울러 가려진 영역을 검출할 수 있다. 가려진 영역의 시차는 가려짐 패턴을 이용하여 올바른 시차를 지정할 수 있어, 결과적으로 보다 정확한 중간 시점 영상을 생성할 수 있다.

Abstract

In this paper, we describe a novel method to generate an intermediate scene using BDM (Bidirectional Disparity Morphing) from the parallel stereopair. Because an image is composed of several layers and each layer has a similar disparity, it is available to use the block based disparity estimation. In order to prevent the false correspondence, however, we closely investigate the corresponding block as we adaptively vary the block size according to the estimation error. Therefore, we can detect the occlusion because of larger estimation error of the occluded region. We define three occluding patterns, which are derived from the peculiar property of the disparity map, in order to smooth the computed disparity map. The filtered disparity map using these patterns presents that the false disparities are well corrected and the boundary between foreground and background becomes sharper. As a result, we can improve the quality of the intermediate scenes.

Key Word : bidirectional disparity morphing, view interpolation, occlusion pattern, 3D TV

* 正會員, 中央大學校 尖端映像大學院 映像工學科
(Dept. of Image Engr., Graduate School of Advanced Imaging Science, Multimedia, and Film, Chung-Ang University)

** 正會員, 電子部品研究院 뉴미디어통신연구센터

(Advanced Media Communication Research Center, Korea Electronics Technology Institute(KETI))

※ 본 연구는 교육부의 두뇌한국21 사업(BK21) 및 과학 기술부의 국가지정 연구실(2000-N-NL-01-C-285) 지원으로 수행되었습니다.

接受日字:2001年1月7日, 수정완료일:2002年3月5日

I. 서 론

3차원 영상 재생기술의 발달과 함께 기존의 평면 영상이 아닌 3차원의 공간적인 느낌을 줄 수 있는 영상 콘텐츠의 개발이 중요하다. 최근의 다시점 재생 장비를 이용하면 제한된 시점이 아닌 변화된 위치에서도 특수한 안경을 착용하지 않고 3차원 공간감을 느낄 수 있다.

보다 자연스럽고 실감적인 3차원 영상을 생성하기 위해서는 많은 시점 영상이 필요하고, 이는 곧 카메라의 증가를 요구한다. 따라서 제한된 카메라의 개수로부터 충분한 시점 영상을 제공하기 위해서는 카메라 사이의 중간 시점 영상을 보간하는 기술이 필요하다.

기존의 중간 영상 보간 기술은 주어진 2차원 영상을 3차원 모델로 재구성한 뒤, 각각의 3차원 좌표를 다시 2차원 영상 평면으로 재투영시켜 새로운 시점 영상을 생성했다. 그러나 이 방법은 3차원 모델링을 하기 위한 제약 조건이 많았으며, 또한 재구성된 많은 데이터를 저장하고 처리하는데 필요한 별도의 하드웨어가 요구되었다^[1-4].

이러한 단점을 극복하기 위한 방법이 2차원에서 2차원으로 직접 변환시키면서 3차원적인 효과를 주기 위한 영상 기반 렌더링(IBR, Image Based Rendering)이 도입되었다. 이 방법은 주어진 두 영상에 대해서 대응되는 두 점 사이의 좌표를 보간하는 방식으로 사전에 주어진 영상에서 각 화소간의 대응점을 알아야 한다.

Levoy와 Hanrahan^[12] 그리고 Gortler^[13] 등은 3차원 표면을 지나는 광선(light rays)을 표현하는 4차원의 필드(field)를 구성하기 위해서 IBR의 개념을 확장시켰다. 새로운 시점 보간 영상은 광선들의 필드가 정확하게 교차하는 2차원 단면을 계산함으로써 생성할 수 있다. 그러나 이 방법의 문제점은 실감나는 새로운 시점 영상을 보간하기 위해서는 무수히 많은 영상이 필요하다. 따라서 실제 장면 모델링에는 부적합하다. Chen^[2]은 두 영상에서 서로 대응되는 점을 찾기 위해서 카메라의 위치, 방향 정보와 해당 장면의 거리 영상(range data)을 함께 이용하여 효율적으로 대응점을 검출하였다. 그러나 이 방법은 대응점 검출을 위해 필요한 사전 정보가 필요하고, 대응되는 화소를 단순히 선형 보간하였기 때문에 보간된 영상이 열화되는 고스트(ghost) 현상이 나타난다. Seitz^[3]는 보간된 영상에서 고스트 현상을 최소화 시키는 뷰 볼핑(view morphing)을 제안하였다.

그러나 이 방법은 주어진 두 영상에서 서로 대응되는 특징점을 정합하기 위하여 사용자가 중간 시점 영상 보간에 직접 참여하기 때문에 자동적인 영상의 재구성이 불가능한 단점을 갖는다. 또한 사용된 영상에 가려짐(occlusion)이 없다고 가정하였기 때문에 일반적인 자연 영상에 적용이 어렵다. Satoh^[4]는 중간 시점 영상을 자동으로 보간하기 위해서 3×3 카메라 행렬을 이용하였다. 이는 다중 스테레오(multiple-based stereo) 방식과 유사한 방법으로 주어진 영상에 대한 시차 맵(disparity map)을 계산한다. 그리고 가려짐 영역에 대한 시차 계산을 위하여 각각 카메라에 대한 가려짐 패턴을 정의하여 배경과 전경의 경계가 보다 명확한 시차 맵을 얻을 수 있었다. 그러나 이 방법은 9개의 카메라를 이용하기 때문에 시차 맵을 얻기 위한 계산량이 크게 증가하는 단점을 갖는다.

본 논문에서 제안된 방법은 그림 1과 같이 서로 평행하게 놓여진 두 대의 카메라로부터 취득된 두 영상에서 발생하는 시차 정보를 이용하여 자동으로 중간 시점에서의 영상을 보간하는 방법을 제안한다. 정확한 중간 시점 영상을 재구성하기 위해서는 조밀한 시차 정보가 필요하지만, 이러한 시차 맵의 계산은 많은 시간이 필요하고 가려짐으로 인한 대응점 검색의 실패는 재구성된 영상의 질을 저하시킨다.

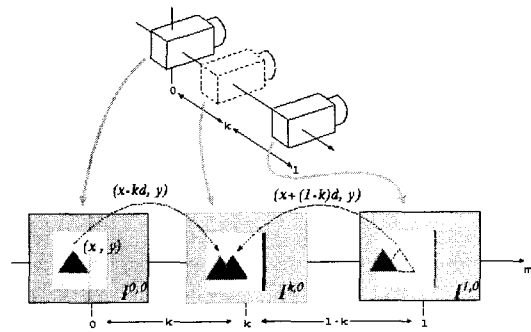


그림 1. 시차 맵을 이용한 중간 시점 영상 보간 예^[17]
Fig. 1. Example of intermediate scene interpolation by using disparity map.

본 논문에서는 영상이 몇 개의 계층(layer)으로 구성되어 있고, 각각의 계층은 유사한 시차를 갖는 성질을 이용하여 블록 단위의 시차 맵을 계산한다. 또한 서로 다른 계층이 겹쳐지는 영역에서는 시차값의 크기가 급격하게 변화하고^[1, 5, 8], 정합 오차가 큰 성질을 이용하여 블록의 크기를 적응적으로 변화시켜 보다 정밀한 시차

맵을 계산한다. 가려짐이 발생하는 영역은 그렇지 않은 영역에 비해서 정합 오차가 크기 때문에, 이 결과로부터 가려짐이 발생한 영역을 검출할 수 있다.

가려짐이 발생한 블록은 시차가 불연속적이고, 해당 블록의 시차는 가려짐이 발생한 블록의 적절한 방향에서 가려지지 않은 블록이 포함된 계층의 시차와 유사한 값을 갖는다^[1]. 우리는 이 성질을 이용하여 가려짐 패턴을 정의할 수 있고, 이 패턴을 이용하여 가려짐이 발생한 블록의 시차를 정확하게 보정하면서, 가려진 영역의 경계를 보다 뚜렷이 할 수 있다.

좌우 두 영상으로부터 가려짐을 고려한 각각의 시차 맵이 계산되면, 좌우 두 영상은 각각의 시차 맵을 이용하여 중간 시점 영상을 재구성할 수 있다. 이렇게 재구성된 두 영상을 디졸브(dissolve) 시킴으로써 시차 계산의 오류를 최소화하는 중간 시점 영상이 생성된다^[1].

본 논문의 구성은 서론에 이어 2장에서 중간 시점 좌표 계산을 위한 양방향 시차 몰핑(BDM)을 설명한다. 3장에서는 적응적 블록 정합 알고리즘을 이용하여 가려짐 영역을 검출하고, 가려짐 영역의 패턴을 정의하여 해당 영역의 시차값을 보정한다. 4장에서는 제안된 방법을 다양한 영상에 적용한 결과와 그 성능을 평가하고, 5장에서 결론을 맺는다.

II. 양방향 시차 몰핑(Bidirectional Disparity Morphing, BDM)

우리는 영상 취득을 위한 카메라가 수평 방향으로만 떨어져 있고, 서로 평행한 카메라 기하 모델로 설정되었다고 가정한다. 공간상의 한 점은 좌우 영상에서 한 쌍의 대응점 P_l, P_r 을 형성한다. 만일 이 두 점이 두 영상에서 모두 보인다면 대응점 사이의 시차는 화소(pixel) 단위의 거리로 정의된다. 우리는 평행한 카메라 기하 모델을 가정하였기 때문에 시차는 수평 방향 성분만이 존재하고, 그 관계식은 식(1)과 같다.

$$\begin{aligned}
 p_r &= \begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} x_l + d_{L \rightarrow R}(x_l, y_l) \\ y_l \end{bmatrix} \\
 &= p_l + \begin{bmatrix} d_{L \rightarrow R}(x_l, y_l) \\ 0 \end{bmatrix} \quad (1)
 \end{aligned}$$

$d_{L \rightarrow R}(x_l, y_l)$ 은 왼쪽 화소에서 대응하는 오른쪽 화소까지의 시차이다.

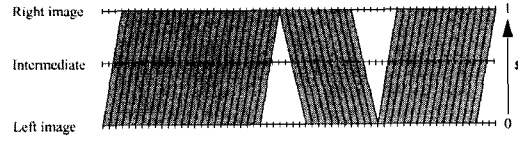


그림 2. 1차원상에서의 샘플 장면^[1]
Fig. 2. Sample scene in 1D.

그림 2는 주어진 스테레오 영상에서 하나의 스캔 라인(scan line) 상에서 대응되는 화소 P_l, P_r 간의 기하학적인 관계와 중간 시점 좌표를 몰핑 계수(morphing coefficient) s 를 통해 유도되는 과정을 보인다. P_l, P_r 간의 거리를 1로 정규화 시키면, 중간 시점 영상은 이 거리의 상대적인 거리($0 < s < 1$)에 따라서 결정된다. 주어진 왼쪽 영상에 대한 중간 시점 영상은 식(2)와 같이 시차의 함수로 표현되며, 오른쪽 영상에 대해서도 같은 방법이 적용된다. 한 쌍의 스테레오 영상은 서로 수평 방향으로만 떨어져 있다고 가정하였기 때문에 y 축 좌표는 변하지 않는다. $s = 0$ 에서 보간된 영상은 왼쪽 영상이고, $s = 1$ 은 오른쪽 영상과 동일하다.

$$(0 < s < 1) \quad (2)$$

식(2)를 이용하여 보간된 영상에는 겹쳐짐(fold)과 홀(hole)이 발생한다^[1-3]. 겹쳐짐은 원 영상에서 보이는 화소가 보간된 영상에서 보이지 않는 경우, 보간된 한 화소에 여러 개의 화소가 겹쳐지게 된다. 그러나 겹쳐지는 화소 중 최종적으로 영상에서 나타나는 화소는 관찰자에게 가장 가까운 곳에 존재하는 물체이고, 관찰자에게 가까운 물체일수록 시차값이 크기 때문에 z-buffer 알고리즘을 이용하면 겹쳐짐에 대한 문제를 해결할 수 있다^[2].

홀은 원 영상에서 보이지 않던 화소가 보간된 영상에서 나타나는 경우에 발생한다. 일반적으로 홀은 주변 화소를 가지고 보간을 하지만, 두 영상간의 시차가 큰 경우에는 주변 화소를 이용한 보간에는 한계가 있다. 따라서 본 논문에서는 MPEG 영상에서 사용하는 움직임 보상 영상(motion compensated image)의 보간 방법을 이용한다. 즉, 스테레오 영상에서 한 쪽 영상에서 보이지 않는 영역은 다른 쪽 영상에서는 보이는 영역이다. 따라서 좌우 각각의 시차 맵으로부터 보간된 영상에 대해서 한 쪽 영상에서 발생한 홀은 다른 쪽 영상에서는 화소값을 갖는 보이는 화소에 대응된다. 식(3)은 최종적으로 디졸브된 영상에서 각 화소의 화소값을 정의

한다. 마지막 단계에서 두 영상을 디졸브 시키는 작업이 부수적일 수 있으나, 두 영상을 디졸브 시킴으로써 시차 맵의 계산에서 발생할 수 있는 오류를 최소화시킬 수 있다.

$$IP_s(x, y) = \begin{cases} (1-s) \cdot IP_{s,L \rightarrow R}(x, y) + s \cdot IP_{s,R \rightarrow L}(x, y) \\ IP_{s,L \rightarrow R}(x, y) \\ IP_{s,R \rightarrow L}(x, y) \\ \text{Nearest neighbor pixel} \end{cases}$$

$$\begin{aligned} & \text{if } IP_{s,L \rightarrow R} \neq 0, IP_{s,R \rightarrow L} \neq 0 \\ & \text{if } IP_{s,L \rightarrow R} \neq 0, IP_{s,R \rightarrow L} = 0 \\ & \text{if } IP_{s,L \rightarrow R} = 0, IP_{s,R \rightarrow L} \neq 0 \\ & \text{if } IP_{s,L \rightarrow R} = 0, IP_{s,R \rightarrow L} = 0 \end{aligned} \quad (3)$$

$IP_s(x, y)$ 는 (x, y) 위치에서의 디졸브된 화소값이고, $IP_{s,L \rightarrow R}(x, y)$, $IP_{s,R \rightarrow L}(x, y)$ 는 각각 좌우 두 영상으로부터 보간된 화소값이다. 좌우 영상으로부터 보간된 두 영상의 화소가 모두 0이 아닌 경우에는 물평 계수(s)에 의해서 평균을 취하고, 두 영상 중 하나의 화소값이 0인 경우 0이 아닌 화소값으로 대체한다. 그리고 두개의 화소값이 모두 0인 경우는 주변의 가장 가까운 화소값으로 보간한다.

III. 적응적 블록 단위의 시차 맵 계산

1. 시차 맵의 계산 및 가려짐 영역 검출

그림 3에서 보여주듯이 영상은 여러 개의 계층으로 구성되어 있고, 각각의 계층은 서로 유사한 시차를 갖고 있기 때문에 본 논문에서는 영상을 8×8 블록으로 나누어 블록 단위의 시차를 계산한다. 이 방법은 영상의 해상도를 줄일 수 있기 때문에 계산량이 크게 감소한다. 그러나 서로 다른 계층이 겹쳐지는 블록에 대해서도 동일한 시차를 부여하기 때문에 합성된 영상의 화질을 떨어뜨리는 원인이 된다. 또한 서로 다른 계층이 겹쳐지는 영역은 일반적으로 다른 영역에 비해서 정합 오차가 크고, 영상의 가려짐이 발생할 확률이 높다. 따라서 본 논문에서는 정합의 결과에 따라서 블록의 크기를 가변적으로 변화시켰다.

기본적인 블록의 크기는 8×8 이고, 정합 오차가 문턱치(th_1) 이상인 영역은 4개의 4×4 블록으로 나누어 보다 세밀한 블록 정합을 수행한다. 이때 가려짐이 발



그림 3. 계층 분리된 영상의 예
Fig. 3. Example of layer image.

생한 영역은 다른 블록에 비해서 여전히 정합 오차가 크다. 따라서 4개의 4×4 블록 중 정합 오차가 문턱치(th_2) 이상인 블록은 영상에서 가려진 영역으로 간주한다. 그림 4는 블록 단위의 시차 맵을 계산하고, 가려짐이 발생하는 영역을 검출하는 알고리즘을 설명한다.

```

Do block matching for 8x8 block
If less than  $th_1$ , do block matching for another 8x8 block
Else
  While do block matching for each 4x4 sub-block
  If less than  $th_2$ , do block matching for another 4x4 block
  Else assign occlusion to this block
  
```

그림 4. 블록 단위 시차 맵의 계산 및 가려짐 블록 검출
Fig. 4. Computation of disparity map and detection of occlusion.

2. 시차 값의 보정

일반적으로 시차값은 연속적이고 동일 계층 내에서 유사하기 때문에, 주변 화소에 대한 시차값의 분포는 균일한 특성을 갖는다. 따라서 시차 맵에서 발생하는 오류에 대한 보정은 주변 화소들의 시차를 평균하는 평균값 필터(mean filter) 보다는 중간값을 취하는 미디언 필터(median filter)의 성능이 우수하다.

가려짐 영역에 대한 시차값의 보정은 미디언 필터만으로는 불가능하다. 가장 보편적인 방법은 Bayesian 이론과 같은 확률과 통계적인 방법을 이용하는 것이다. 그러나 이 방법은 다루기가 어렵고, 계산량이 많기 때문에 실시간 구현이 쉽지 않다.

McVeigh와 Satoh는 가려짐이 발생하는 몇가지의 경우를 가정하여 시차 값을 보정하였다. McVeigh^[1]는 가려짐이 발생한 영역의 시차값은 가려진 영역의 적절한 위치(appropriate side)에 존재하는 가려지지 않은 화소의 시차와 동일하다고 여겼다. 즉, 가려진 영역이 왼쪽 영상에서 보인다면 적절한 위치는 가려짐이 발생한 영역의 왼쪽 방향에 있다는 것이다

Satoh^[5]는 가려짐 영역의 경계는 선형적이라고 가정

하여 9개의 가려짐 패턴을 정의하였다. 이 패턴은 등간격으로 구성된 3×3 카메라 행렬에 적용하여 중앙 영상에 대한 주변 8개의 영상과의 대응점을 검출하였다. 이 경우, 가려짐이 발생한 영상에 대한 대응점 검출은 생략하기 때문에 대응점 검출 실패는 발생하지 않는다. 이 방법은 좋은 결과를 나타내고 있으나, 주변 8개 영상에 대해서 9개의 패턴을 모두 적용해야 하기 때문에 계산량이 증가한다. 그 결과 실시간으로 시차 값을 계산하여 응용하는 것은 불가능하다.

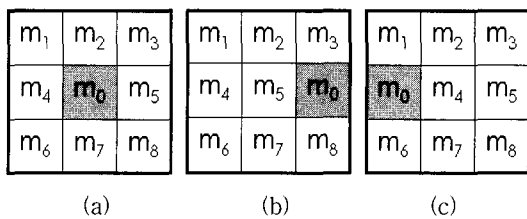


그림 5. 가려짐 패턴 (a) 가려짐이 발생하지 않은 화소의 패턴 (b) 왼쪽 영상에서 가려짐이 발생한 화소의 패턴 (c) 오른쪽 영상에서 가려짐이 발생한 화소의 패턴

Fig. 5. Occluding patterns. (a) pattern of the both visible points. (b) pattern of the left occluded points. (c) pattern of the right occluded points.

본 논문에서는 잘못 지정된 시차와 가려짐 영역의 시차를 보정하기 위해서 3×3 미디언 필터를 이용한다. 기본적인 카메라의 설정이 서로 수평으로 떨어져 있고, 카메라는 평행하다고 가정하였기 때문에 가려짐이 발생하는 경계면은 수직적인 경향을 나타낸다. 또한 동일 계층 내의 시차값은 유사하다. 이 두 성질을 이용하면 그림 5와 같은 블록 단위의 가려짐 패턴(occluding pattern)을 정의할 수 있다. 그림 5(a)는 가려짐이 발생하지 않은 블록에 대한 일반적인 미디언 필터 마스크를 나타낸다. 왼쪽 영상에서 존재하는 가려짐 영역은 전경 물체가 왼쪽으로 이동하면서 발생하고, 가려짐의 경계는 전경 물체의 왼쪽에서 나타난다. 오른쪽 영상에서 존재하는 가려짐 영역은 이와 반대의 성질을 갖는다. 이러한 특성은 그림 5(b), (c)와 같이 중심 화소를 가려짐이 발생하는 영역의 경계선 상에 놓음으로써 가려짐 경계를 뚜렷하게 하고, 정확한 시차를 지정할 수 있다.

IV. 모의 실험 및 토의

본 논문에서 제안된 방법은 그림 6의 세 가지의 영상에 적용하였다. 그림 6(a)는 320×240의 RGB 컬러 영상이고, 그림 6(b)는 320×240, 그림 6(c)는 240×256의 그레이 영상이다. 그림 6(b), (c)는 카메라를 수평 방향으로 0.5인치(inch)씩 이동시켜서 취득한 영상으로 제안된 알고리즘의 성능을 평가할 수 있는 실험 영상이다. 제안된 알고리즘은 화소 기반의 정합 방법과 고정된 블록 크기를 갖는 블록 정합 방법을 이용한 보간 방법과 비교된다. 화소 단위의 시차 맵을 계산하기 위해서 15×15 블록을 이용하였고, 블록 단위의 시차 맵을 계산

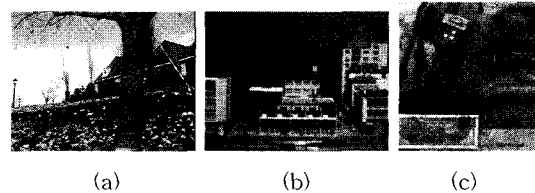


그림 6. 실험 영상 (a) Flower garden, (b) Building, (c) Coal
Fig. 6. A set of experimented images. (a) Flower garden, (b) Building, (c) Coal

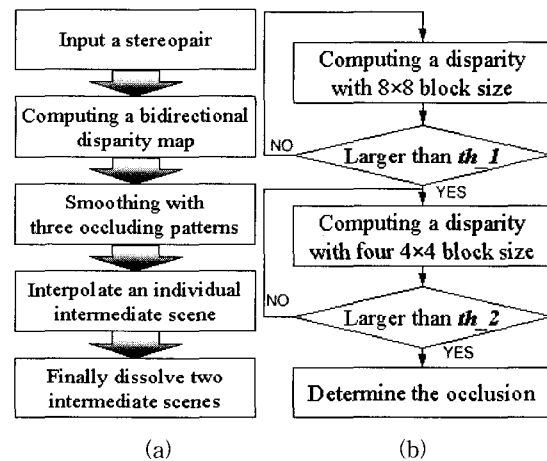


그림 7. 제안된 알고리즘의 흐름도 (a) 전체적인 알고리즘의 흐름도 (b) 시차 맵을 계산하기 위한 적응적 블록 정합 알고리즘의 흐름도
Fig. 7. Flow chart of the proposed algorithm. (a) overall flow chart of the proposed algorithm, (b) adaptive block matching algorithm for computing a disparity map.

하기 위해서 8×8 블록과 4×4 블록을 이용하였다. 대응점 검출을 위한 검색 영역은 $-15 \leq d_x \leq 15$ 이다.

그림 7은 제안된 알고리즘의 전체적인 흐름도를 나타내고 있다. 그림 7(a)에서 입력된 한 쌍의 스테레오 영상은 좌우 각 영상에 대해서 시차 맵을 계산한다. 시차 맵의 계산은 그림 7(b)와 같은 과정을 거쳐 서로 다른 계층이 겹쳐지는 경계면에서의 시차를 보다 정확하게 계산하면서 가려진 영역을 검출한다. 계산된 시차 맵에 존재하는 오류를 제거하고, 가려진 영역에 올바른 시차를 지정하기 위해서 세가지 가려진 패턴을 이용하여 스무딩을 수행한다. 스무딩을 마친 시차 맵을 이용하여 좌우 영상은 각각 원하는 중간 시점 영상으로 보간되고, 보간된 두 영상은 시차 맵에 존재하는 오류를 최소화하기 위해 서로 디졸브 시켜 최종적인 중간 시점 영상으로 재구성된다.

1. 가려짐 영역의 검출

주어진 영상에 대한 시차를 계산하기 위해서 식(4)와 같이 MAD(Mean Absolute Difference)를 평가함수로 이용하였다. 이 평가함수는 대응되는 블록이 서로 정확하게 정합하면 할수록 작은 값을 나타내기 때문에 검색 영역 안에서 최소값을 나타내는 화소를 대응점으로 결정하였다. 그러나 서로 다른 계층이 겹쳐져서 가려짐이 발생한 영역은 대응점이 존재하지 않기 때문에 단순히 최소의 평가함수에서 대응점을 선택하는 것은 불합리하다. 따라서 대응점 결정의 평가 기준이 필요하다.

$$E(d_x) = \sum_{(x,y) \in B} |I_m(x,y) - I_{m+1}(x+d_x,y)|$$

$$[\hat{d}_x] = \arg \min_{(d_x)} [E(d_x)] \quad (4)$$

본 논문에서는 반복적인 실험을 통하여 블록의 최대 오차값의 10%를 문턱치(th_1)로 결정하여 그림 7(b)와 같이 대응점 검색의 결과를 평가하였다. 즉, th_1 보다 작은 값을 나타내는 영역은 대응점의 검색이 성공하였다고 간주하고 그렇지 않은 영역은 실패하였다고 간주하였다. th_1 보다 큰 값을 나타내는 영역은 4개의 4×4 블록으로 나누어서 보다 세밀한 검색을 수행한다. 이 경우 역시 해당 블록의 최대 오차값의 10%를 문턱치(th_2)로 결정한다. 4×4 블록을 이용하는 경우에도 가려짐을 나타내는 영역은 여전히 그렇지 않은 영역에 비해서 오차값이 크다. 따라서 th_2 보다 큰 값을 나타내는 영역은 가려짐이 발생한 영역으로 간주한다.

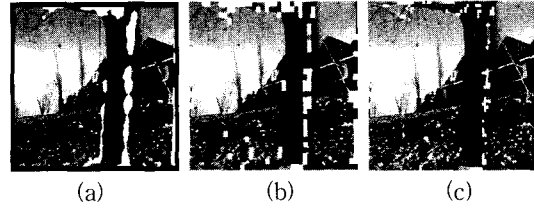


그림 8. 검출된 가려짐 영역 (a) 화소 기반의 정합 방법 (b) 고정된 크기의 블록 정합 방법 (c) 적응적 블록 정합 방법

Fig. 8. Detection of occlusion. (a) Pixel based method, (b) Fixed block matching method, (c) Adaptive block matching method.

그림 8은 그림 6(a)에서 가려진 영역을 검출한 결과를 나타내고 있다. 그림 8(a)는 15×15 블록을 이용하여 화소 기반의 시차 맵을 계산할 때 검출된 가려짐 영역이고, 그림 8(b)는 고정된 8×8 블록에 대한 결과이다. 두 결과는 서로 비슷한 결과를 보이고 있으며, 블록의 크기가 고정되어 있기 때문에 실제 가려지지 않은 영역까지 가려진 영역으로 검출하고 있다. 이와 대조적으로 그림 8(c)는 8×8 과 4×4 의 두 블록을 적응적으로 변화시키면서 가려진 영역을 검출한 결과, 가려짐이 발생하는 나무의 경계면을 보다 세밀하게 검출하였다.

2. 가려짐 패턴을 이용한 스무딩(smoothing)

위에서 얻어진 시차 맵은 신뢰성을 높이기 위해서 3.2절에서 정의한 세 개의 가려짐 패턴을 이용하여 스무딩을 수행한다. 두 영상에서 모두 보이는 블록은 그림 5(a)의 마스크를 이용해서 일반적인 미디언 필터링을 수행하고, 왼쪽과 오른쪽 영상에서 검출된 가려진 블록은 그림 5(b)와 그림 5(c)를 이용해서 각각 스무딩을 수행한다. 그림 9는 그림 6(a)에 대한 각각의 시차 맵 계산 방법들에 대한 스무딩 결과를 보여준다. 그림에서 보여 주듯이, 스무딩의 결과 가려짐이 발생하는 나무의 경계면이 보다 뚜렷해졌음을 확인할 수 있다. 또한 그림 9(a), (c)를 비교해 보면 적응적 블록 정합으로 얻어진 시차 맵은 화소 단위의 시차 맵과 유사하다.

3. 영상의 합성

좌우 두 영상으로부터 각각의 시차 맵이 계산이 되었다면, 각각의 영상은 중간 시점 영상으로 보간할 수 있다. 이렇게 보간된 두 영상은 시차 맵의 계산에서 발생하는 오류를 최소화하기 위해서 디졸브 시킨다. 대응 화소의 화소값은 식(3)을 이용해서 화소값을 지정한다. 그림 10은 그림 6의 실험 영상에 대해서 물핑 계수를

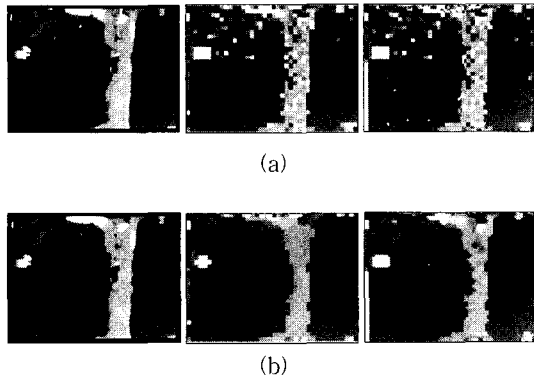


그림 9. 스무딩 적용 전, 후의 시차 맵 (a) 스무딩을 적용하기 전의 시차 맵 (b) 스무딩을 적용한 후의 시차 맵

Fig. 9. Disparity maps before and after smoothing. (a) before smoothing, (b) after smoothing.

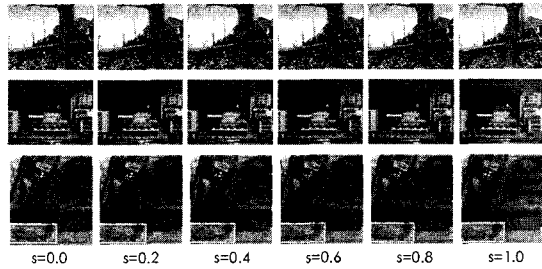


그림 10. 중간 시점 보간 영상의 예
Fig. 10. Examples of intermediate images.

변화시키면서 보간된 중간 시점 영상의 결과를 나타내고 있다.

4. 성능평가

제안된 알고리즘은 화소 기반의 방법과 고정된 블록 크기의 블록 정합 방법과 비교된다. 그림 6(b), (c)는 카메라의 간격을 0.5인치씩 수평 방향으로 이동시켜 취

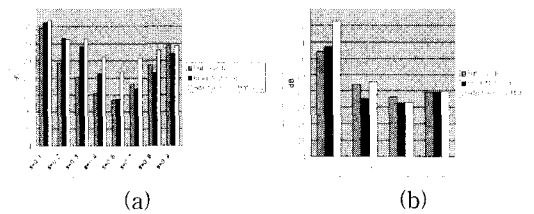


그림 11. PSNR 결과 (a) Building image (b) Coal image
Fig. 11. Results of PSNR. (a) Building image, (b) Coal image.



(a)



(b)



(c)

그림 12. 각각의 보간법을 이용하여 생성된 중간 시점 영상 예($s=0.5$). (a) 화소기반 블록 정합 알고리즘, (b) 일반적인 블록 정합 알고리즘, (c) 적응적 블록 정합 알고리즘

Fig. 12. Examples of the intermediate scene generated by each interpolation method. (a) pixel based block matching algorithm, (b) normal block matching algorithm, (c) adaptive block matching algorithm.

특하였기 때문에 격 간격으로 선택한 두 영상의 $s = 0.5$ 에서 보간된 영상은 중앙의 원 영상과 비교가 가능하다. 각각의 방법에 대한 성능을 평가하기 위해서 식(5)의 PSNR을 이용한다.

$$MSE = \frac{1}{MN} \sum_{y=1}^M \sum_{x=1}^N [I(x, y) - I'(x, y)]^2$$

$$PSNR = 20 \times \log_{10} \left(\frac{255}{\sqrt{MSE}} \right) \quad (5)$$

M, N 은 각각 영상의 가로, 세로 크기이고 $I(\cdot)$ 와 $I'(\cdot)$ 은 원 영상과 시점 보간된 영상을 나타낸다.

그림 11은 그림 6(b), (c)에 대한 PSNR 결과를 나타내고 있다. 일반적으로 화소 기반의 시차 맵이 가장 정확하다고 여겨지고 있으나, 대부분의 보간된 영상에 대해서 제안된 방법의 성능이 1~2dB 높거나 유사한 성능을 보이고 있다. 이는 화소 단위의 방법이 고정된 크기의 블록을 이용하는 반면에 제안된 방법은 블록의 크기를 적응적으로 변화시켰기 때문에 가려짐 영역을 보다 정확하게 검출할 수 있었다. 이러한 가려짐 영역의 정확한 검출은 보간된 영상의 질을 향상시키는 원인이 된다.

그림 12는 기존의 방법과 제안된 방법을 각각 이용하여 보간된 영상을 보여주고 있다. 그림 12(a)는 전경과 배경간의 경계 뿐만 아니라 좌측의 가로등과 우측의 전경 나무와 배경 집 사이의 나무 가지에서도 역시 많은 열화가 존재한다. 이에 반해, 그림 12(c)의 경우 영상의 열화가 많이 제거 되었으며, 특히 전경과 배경의 경계가 뚜렷해져 전체적으로 영상을 개선시켰다.

표 1은 중간 시점 영상을 보간하기 위해 필요한 계산 시간을 나타내고 있다. 화소기반의 방법은 40초 이상의 시간이 소요되는 반면, 8×8 블록 기반의 방법은 0.5초가 소요되어 계산시간이 약 1/80로 줄었다. 적응적 블록 기반의 방법은 가려짐 영역에서는 4개의 4×4 블

표 1. 중간 시점 영상 보간에 필요한 처리 시간(단위 : 초)

Table 1. Time table for interpolating an intermediate scene(unit : second).

| Images Methods | Flower garden | Building | Coal |
|-------------------------|---------------|----------|-------|
| Pixel based Method | 43.96 | 42.88 | 35.36 |
| Fixed Block Matching | 0.5 | 0.5 | 0.41 |
| Adaptive Block Matching | 0.55 | 0.52 | 0.41 |

록으로 나누어져 세밀한 작업을 수행하기 때문에 약간의 시간(약 0.1초 미만)이 더 소모된다. 그러나 가려짐 영역이 없을 경우는 블록 기반의 방법과 동일한 계산 시간이 요구된다.

V. 결 론

본 논문은 평행하게 놓인 두 대의 카메라로부터 취득된 한 쌍의 스테레오 영상에 대해서 양방향 시차 맵을 구성하고, 이 시차 정보를 이용하여 중간 시점 영상을 보간하는 알고리즘을 제안하였다. 영상은 하나 이상의 계층으로 구성되었고, 각각의 계층은 유사한 시차를 갖기 때문에 블록 기반의 정합 방법은 시차 맵을 구성하기에 적합하다. 그리고 서로 다른 계층이 겹쳐지는 영역은 그렇지 않은 영역에 비해서 오차가 크고, 가려짐이 발생할 확률이 높기 때문에 블록의 크기를 적응적으로 변화시킴으로써 겹쳐진 영역에서 보다 정확한 대응점을 검출하였고, 가려진 영역 또한 검출할 수 있었다. 계산된 시차 맵의 오류를 보정하고, 가려진 영역의 시차를 정확하게 지정하기 위해서 가려짐 패턴을 정의하였다. 이 패턴은 카메라 이동에 따른 가려짐이 발생하는 방향성과 가려짐 영역의 경계를 고려하여 일반적인 미디언 필터의 중심 화소의 위치를 변화시켰다. 그 결과, 가려진 영역에 정확한 시차를 지정할 수 있었고, 가려진 영역의 경계를 보다 뚜렷이 할 수 있었다. 제안된 방법을 여러 실험영상에 적용하여 그 성능을 비교한 결과, 제안된 방법에서 영상의 PSNR이 증가하였고, 계산에 필요한 시간을 약 1/80로 줄일 수 있었다.

참 고 문 헌

- [1] J. S. McVeigh, M. W. Siegel, and A. G. Jordan, "Intermediate view synthesis considering occluded and ambiguously referenced image regions," *Signal Processing : Image Communication*, Vol. 9, pp. 21~28, 1996.
- [2] S. E. Chen, L. Williams, "View Interpolation for Image Synthesis," In *Proc. SIGGRAPH 93*, pp. 279~288, 1993.
- [3] S. M. Seitz, C. R. Dyer, "Physically-valid view synthesis by image interpolation," In *Proc. IEEE Workshop on Representation of Visual*

Scenes, pp. 18~25, 1995.

[4] K. Satoh, I. Kitahara, Y. Ohta, "3D Image Display with Motion Parallax by Camera Matrix Stereo," Proc. of Third Int. Conf. on Multimedia Computing and Systems(ICMCS'96), 1996.

[5] K. Satoh, Y. Ohta, "Occlusion Detectable Stereo Using A Camera Matrix," Proceedings Second Asian Conference on Computer Vision(ACCV'95), Vol.2, pp. 331~335, 1995.

[6] S. M. Seitz, C. R. Dyer, "View Morphing," Proc. Of SIGGRAPH'96, pp. 21~30, 1996.

[7] C. Cafforio, F. Rocca, S. Tubaro, "Motion Compensated Image Interpolation," IEEE Trans. On Communications, Vol.38, No.2, pp. 215~222, Feb. 1990.

[8] C. L. Zitnicj, T. Kanade, "A Cooperative Algorithm for Stereo Matching and Occlusion Detection," IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol.22, No.7, pp. 675~684, July 2000.

[9] 김대현, 최중수, "Bidirectional disparity morphing 을 이용한 view interpolation", 대한전자 공학회 하계합동학술대회 논문집, 2001년 6월

[10] A. Redert, E. Hendriks, J. Biemond, "3D Scene Reconstruction with Viewpoint Adaptation on Stereo Displays," IEEE Trans. On Circuits and Systems for Video Technology, Vol.10, No.4, pp. 550~562, June 2000.

[11] J. R. Ohm, E. Izquierdo M., "An Object-Based System for Stereoscopic Viewpoint Synthesis," IEEE Trans. On Circuits and Systems for Video Technology, Vol.5, No.5, pp. 801~811, Oct. 2000.

[12] M. Levoy and P. Hanrahan, "Light Field Rendering", In Proc. of SIGGRAPH'96, 1996.

[13] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen, "The Lumigraph", In Proc. of SIGGRAPH'96, 1996.

[14] Y. Sugaya, Y. Ohta, "Stereo by integration of two algorithms with/without occlusion handling," Proc. 15th International Conference on Pattern Recognition(ICPR2000), Vol.1, pp. 109~113, Sep. 2000.

[15] M. A. Akhloufi, V. Polotski, P. Cohen, "Virtual View Synthesis from Uncalibrated Stereo Cameras," Proc. Of the IEEE Multimedia Systems, Vol.2, June 1999.

[16] M. Lhuillier, L. Quan, "Image Interpolation by Joint View Triangulation," Proc. Of Computer Vision and Pattern Recognition(CVPR'95), 1995.

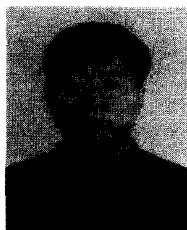
[17] <http://image-gw.esys.tsukuba.ac.jp/research/3DTV/3dtv.html>.

저 자 소 개

金 大 顯(學生會員) 第38卷 SP編 第4號 參照

尹 溶 仁(學生會員) 第38卷 SP編 第4號 參照

崔 宗 秀(正會員) 第38卷 SP編 第4號 參照



金 帝 佑(正會員)
1997년 2월 : 서울시립대학교 제어계측공학과 졸업(공학사). 1999년 2월 : 서울시립대학교 제어계측공학과 졸업(공학석사). 1999년 1월~현재 : 전자부품연구원 전임연구원. <주관심분야 : 영상 압축 알고리즘, 영상통신시스템 관련 알고리즘, 3차원 영상통신 알고리즘>



崔 炳 胡(正會員)
1991년 2월 : 한양대학교 전자공학과 졸업(공학사). 1993년 8월 : 한양대학교 전자공학과 졸업(공학석사). 1993년 7월~1997년 8월 : LG 전자 Video 연구소 1997년 9월~현재 : 전자부품연구원 선임연구원.

<주관심분야 : 영상통신, 영상압축, 3차원 신호처리, ASIC 설계>