

다해상도 기반 계층적 변이맵을 이용한 중간영상 생성 방법

An Intermediate Image Generation Method using Multiresolution-based Hierarchical Disparity Map

허경무*, 유재민
(Kyung Moo Huh and Jae Min Yu)

Abstract : An intermediate images generation method using multi-resolution based hierarchical block matching disparity map is proposed. This method is composed of a disparity estimation, an occlusion detection and intermediate image synthesis. For the disparity estimation, which is one of the important processes in intermediate image synthesis, we use the multi-resolution based hierarchical block matching algorithm to overcome the imperfect ness of block matching algorithm. The proposed method makes disparity maps more accurate and dense by multi-resolution based hierarchical block matching, and the estimated disparity maps are used to generate intermediate images of stereo images. Generated intermediate images show 0.1~1.4 dB higher PSNR than the images obtained by block matching algorithm.

Keywords : intermediate images, multi-resolution, hierarchical block matching, disparity estimation, PSNR

I. 서론

우리가 살고 있는 세계는 2차원의 세계가 아니고 깊이감까지 포함된 3차원의 세계이다. 인간이 3차원으로 입체감을 느낄 수 있는 원인은 물체의 크기, 물체의 명암, 운동시차, 양안시차 등 여러 가지 요인이 있으나, 그중에서 인간의 좌,우 눈이 약 6.5cm 간격으로 떨어져있는 양안시차를 이용하여 양쪽 눈을 통해 본 각각 다른 영상을 뇌에서 합성하여 입체감을 느끼게 되며, 이것을 스테레오 영상[1]이라고 한다.

스테레오 영상은 2안으로 촬영한 좌우 영상을 이용하기 때문에 영상시점의 변화에 관계없이 정지된 입체감만을 느끼고, 카메라 설정이나 사람에 따라 인식되는 거리감에도 차이가 발생하며, 경우에 따라서 거리감 인식이 불가능하거나 심한 피로감을 느끼게 된다. 이에 대하여 다시점 카메라로 획득한 영상을 이용함으로써 시점을 확장하는 동시에 관찰자가 위치를 이동해도 서로 다른 입체영상을 볼 수 있는 자연스러운 3차원 입체감을 느낄 수 있다. 그러나 다시점 영상도 카메라간의 간격이 일정하게 정해져 있기 때문에 시점을 확장할지라도 자연스러운 입체감을 얻기에 한계가 있다. 따라서, 한정된 시점의 영상으로 자연스럽게 연속적인 영상을 제공하고 시점을 확장하기 위한 기술로 영상을 합성하는 방식이 있다.

본 논문에서는 영상의 시점을 확장하기 위한 방법으로 각 시점사이에 또 다른 시점을 갖는 영상(중간영상)을 생성하는 다음과 같은 중간영상생성 방법을 제안한다. 제안하는 방법은 블록기반 변이 추정방식의 효율성을 높이고 초기 변이값을 이용하고 탐색영역의 한정으로 계산량을 감소시킬 수 있는 다해상도 기반 계층적 블록추정기법을 이용한

중간영상생성 방법을 제안한다. 제안하는 방법은 블록정합을 이용한 중간영상생성 결과보다 더 좋은 중간영상생성 결과와 연산시간에서도 7~10% 감소하는 결과를 얻었다.

II. 스테레오 카메라의 구조

스테레오 카메라는 이름에서도 알 수 있는 것처럼 두개의 카메라가 좌영상과 우영상을 각각 얻기 위해서 같이 붙어 있는 기본적인 구조로 되어있다.

크게 스테레오 카메라는 3가지의 구조로 구분을 한다. 첫째는 그림 1과 같은 평행 카메라(parallel camera configuration) 구조로 두 카메라의 렌즈축이 평행하며, 각 카메라의 렌즈축과 광학축이 일치하는 가장 간단한 구조로, 카메라 사이의 거리가 멀수록 두 카메라 사이의 공통 시역이 매우 좁아지는 단점이 있다. 두 번째는 평행 이동식 카메라 구조로 두 카메라의 렌즈축은 서로 평행하나, 광학축이 서로 달라서 평행식 카메라에 비해서 공통시역을 넓힐 수 있는 장점이 있다. 세 번째는 그림 2와 같은 광학축이 안쪽으로 들어온 (Toed-in approach)구조로 각 카메라의 렌즈축과 광학축을 일치하도록 유지하면서 두 카메라의 광학축을 서로 마주보도록 설계한 카메라 구조이다.

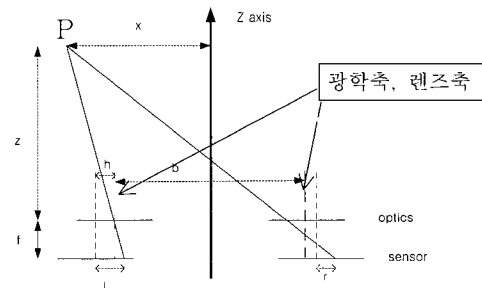


그림 1. 평행 스테레오 카메라의 구조.
Fig. 1. The structure of parallel stereo camera.

* 책임저자(Corresponding Author)

논문접수 : 2003. 7. 10., 채택확정 : 2003. 9. 16.

허경무, 유재민 : 단국대학교 전자컴퓨터학부

(hubkm@dku.edu/you6389@netian.com)

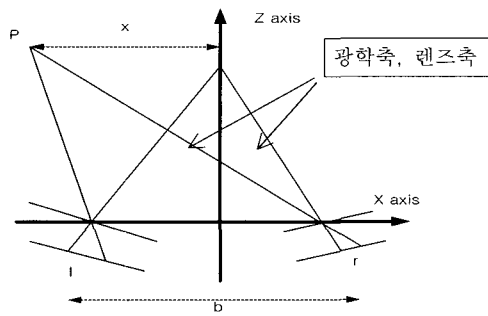


그림 2. Toed-in 스테레오 카메라 구조.
Fig. 2. The structure of Toed-in stereo camera.

이 구조는 공통 시역은 넓으나, 구조의 분석이 어렵다는 단점이 있다.

평행 스테레오 카메라의 최대의 장점은 수직 변이가 없다는 것이다. 본 연구에서는 이 카메라를 기준으로 모든 실험을 했음을 미리 밝혀둔다. 평행 카메라의 경우에 변이를 구해보자. 그림1을 보면 임의의 점 P는 좌·우 카메라 렌즈를 통해 좌·우 영상의 평면에 투영된다. 삼각형의 원리에서부터 아래와 같은 관계가 성립된다. 그림 1에서

$$\frac{x - \frac{b}{2}}{z} = \frac{x - \frac{b}{2} + l - h}{z + f} \tag{1}$$

$$l = \frac{x - \frac{b}{2}}{z} f + h = \frac{fx}{z} - \frac{fb}{2z} + h \tag{2}$$

가 되고, 우측에 대해 같은 방법으로

$$r = \frac{x + \frac{b}{2}}{z} f - h = \frac{fx}{z} + \frac{fb}{2z} - h \tag{3}$$

가 된다. 그래서 변이는

$$d_0 = (l - r) = 2h - \frac{fb}{z} \tag{4}$$

과 같이 구해진다.

III. 중간영상 생성

스테레오 영상은 그 시점이 고정되어 있으므로 안경식으로 화면을 표시하기에는 여러 방향에서 관찰할 수 있는 자유도가 존재하지 않고, 무안경식의 경우에는 단 한사람의 관찰자가 고정된 위치에서만 관찰 가능한 단점이 있으며, 다시점 영상이라 하더라도 데이터량의 증가를 감수해야 하고 시점간의 불연속성이 존재하는 단점이 있다. 이를 해결할 수 있는 방법이 바로 존재하는 영상들사이의 시점 영상을 좌우 영상들로부터 합성해 내는 중간영상 합성이다.

중간영상을 합성하기 위하여 변이추정방법과 가려진 영역을 추출하는 방법을 알아보고, 또한 가려진 영역을 고려하여 다해상도 기반 계층적 블록정합 변이 맵을 이용한 중간영상을 합성하는 방법을 제안한다.

1. 블록정합(block matching algorithm, BMA) 변이추정

변이추정이란 스테레오 영상의 좌,우영상이 공간적 상관성이 매우 높다는 점을 이용하여 서로 대응되는 점 또는 영역을 찾아내는 과정이다. 두 개의 좌,우 이미지를 이용하여 변이를 추정한다함은 간단히 말해서 좌,우 영상이 얼마만큼의 시차변이를 갖는가를 알아내는 것을 의미한다. 두 이미지가 가장 많은 시차변이를 나타내는 부분은 아마도 가까운 부분, 즉 시차에 의한 변화량이 가장 큰 부분일 것이다. 따라서 중간 영상을 합성하는데 있어서 변이추정은 가장 중요한 과정의 하나이며, 이러한 변이를 이용해 중간영상을 예측할 수도 있다. 그렇기 때문에 어떠한 방식으로 변이를 얻어내느냐 하는 것은 중간영상을 얻어내는데 가장 중요한 요소이다.

블록 정합은 왼쪽 영상과 오른쪽 영상을 일정한 크기의 블록단위로 나누고, 기준이 되는 영상의 블록과 다른 쪽 영상에서 한 블록씩 대응해가며 블록당 화소의 변화가 가장 작은 블록을 찾는다. 이때 기준영상의 블록에 비해 얼마나 이동되었는지가 판별된다. 이때 블록의 이동거리로부터 양쪽 이미지에서 느껴지는 디번이가 얼마가 되는지를 알아내어 이를 같은 크기의 이미지 셀에 행렬식으로 이동수치를 적어 넣는 방법이다.

기준이미지의 블록으로부터의 변이량을 추정하기 위한 비용함수로는 다음과 같은 관계를 갖는 MAE (mean absolute error)와 MSE(mean square error)를 사용한다[7].

$$MAE(i, j) = \frac{1}{NM} \sum_{k=1}^N \sum_{l=1}^M | I_r(k, l) - I_l(k + i, l + j) | \tag{5}$$

$$MSE(i, j) = \frac{1}{NM} \sum_{k=1}^N \sum_{l=1}^M [I_r(k, l) - I_l(k + i, l + j)]^2 \tag{6}$$

여기서 N×M은 블록의 크기, I_r은 우영상 그리고 I_l은 좌영상을 나타낸다.

실험에서는 두 이미지의 수직변이가 없는 것으로 가정하고 수평변이만을 고려하여, 블록비교는 각 행별로 독립적으로 비교하였다. 우선 MAE를 통해 각 화소당 변이량의 평균값을 구한 후 블록간의 이동량을 MSE를 사용하여 구하였다. 여기서 기준 이미지의 한 블록과 현재이미지의 모든 블록과의 MSE중에서 가장 작은 MSE를 갖는 블록까지의 이동량을 기준 이미지와 현재 이미지 사이의 수평 변이량(dij)으로 볼 수 있다. 이제 상기의 방법을 통해 각 블록당 원점(0,0)으로부터의 이동거리(화소단위)를 따져서 각 블록에 있는 화소들에 이 벡터를 적용하게 된다.

2. 다해상도 기반 계층적 변이맵을 이용한 중간영상 생성

다해상도 기반 변이추정은 먼저 저역 필터를 거쳐 가우시안 피라미드 영상을 생성한다. 그리고 낮은 해상도에서부터 높은 해상도의 방향으로 변이를 추정한다. 그러므로 이 과정을 다단계 블록 정합 알고리즘이라고도 한다. 이 방법은 블록 정합의 효과성을 개선하기 위하여 제안되었다. 해상도가 높아짐에 따라 변이의 범위가 줄어들므로 계산량이 줄어들게 된다.

2.1 다해상도 영상분해

영상신호의 다해상도 분석은 영상의 해상도를 변환해가며, 필요한 처리를 계층적으로 수행하는 분석 방식이다. 일

반적으로 다해상도 분석은 고정 해상도에서 처리하는 방식보다 연산시간을 단축시킨다는 장점과 인간의 시각 체계가 기본적으로 다해상도 처리를 기반으로 한다는 사실[2]에서 출발한다. 즉 다루어야 할 많은 양의 신호를 분리하여 적은 양의 데이터를 만들기 위하여 원 신호의 대역 제한 필터를 통과시킨 후 신호를 다운 샘플링하여 처리하는 방식을 말한다. 이때 사용되는 대역 제한 필터는 다음과 같은 조건을 만족하여야 하고, 본 논문에서는 가우시안 필터를 대역 제한 필터로 사용하였다[3]. 가우시안 필터는 시간축의 계수들이 가우시안 분포의 모양을 갖기 때문에 붙여진 이름으로 주파수 축의 표현 역시 가우시안으로 나타나기 때문에 모든 조건을 만족한다.

- 조건 1. 필터 커널이 분리 가능해야 한다.
- 조건 2. 시간 축에서의 임펄스 반응과 주파수축의 반응이 모두 회전대칭이어야 한다.
- 조건 3. 필터 커널은 대칭이고 홀수 개의 계수이어야 한다.
- 조건 4. 시간축의 임펄스 반응 및 주파수 축의 반응이 단조 감소의 형태를 가지며, 모두 양수로 이루어져야 한다.

주어진 영상은 원 해상도 영상이 되고, 가우시안 저역통과 필터로 대역을 제한한 후, 가로, 세로축 모두에 대해 1/2 다운 샘플링을 수행하여 화소 수는 1/4이 된다. 이러한 과정을 반복하여 그림 3과 같은 영상 피라미드를 만든다.

2.2 다해상도 기반 계층적 변이추정(multi-resolution based hierarchical block matching algorithm, HBMA)

변이는 블록 정합에 의한 탐색으로 계산된다. 그림 4는 정합과정을 일차원적으로 보여주고 있다.

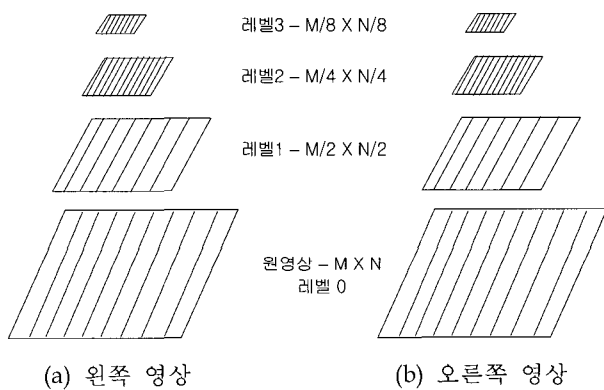


그림 3. 영상의 다해상도 분해.
Fig. 3. Multi-resolution decomposition of images.

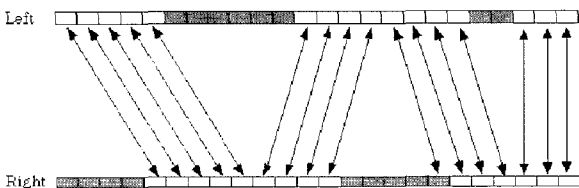


그림 4. 정합과 가려진 영역의 개념.
Fig. 4. Concept of matching and occlusion region.

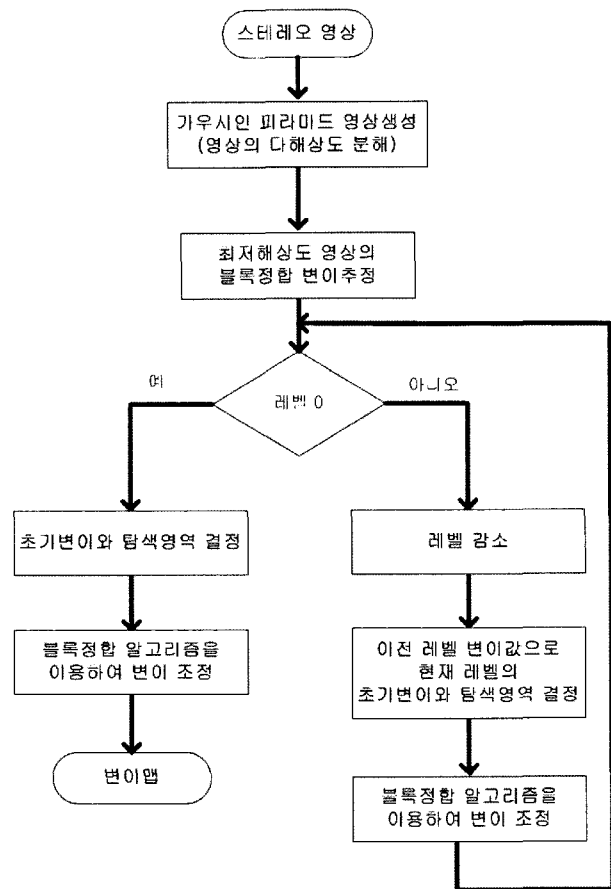


그림 5. 다해상도 기반 계층적 변이추정의 흐름도.
Fig. 5. The flowchart of multi-resolution based hierarchical disparity estimation.

그림 4에서 알 수 있듯이 궁극적인 목표는 좌우 영상간의 대응점을 정확히 찾고, 대응점이 없는 영역(가려진 영역이라고도 함)을 가려내는 것이다. 이것은 물론 대응점만 정확히 찾을 수 있다면 대응되지 않는 부분도 정확하다고 할 수 있지만 대부분의 경우는 대응점을 정확히 찾는 것은 대단히 어려운 일이다. 변이의 정확한 추정은 합성된 중간영상에 증대한 영향을 준다.

다해상도 기반 변이추정 방식은 추정하고자 하는 스테레오 좌, 우 두 영상을 저역 통과 필터와 다운 샘플링 과정을 통하여 사용하여 여러 가지 해상도로 분해한 후 최상위 레벨에서부터 최하위 레벨까지 계층적으로 변이 추정을 하는 방식이다. 낮은 해상도 영상에서 얻어진 변이는 높은 해상도 영상에서 사용된다. 즉 다음 레벨에서 변이 범위의 설정에 사용된다. 해상도가 높아질수록 변이 범위는 줄어든다. 그러므로 계산시간을 단축시킬 수 있다.

그림 5는 다해상도 기반 계층적 변이를 추정하는 과정을 나타낸 그림이다. 먼저 다해상도 분해로 얻어진 좌, 우 스테레오 영상을 통해 변이추정을 하게 되는데, 먼저 최상위 레벨에서 변이 추정을 시작하여 초기 변이 벡터값을 구하게 된다. 이때 계산된 변이 벡터는 그 다음 레벨의 변이 벡터의 초기 값이 되어 이를 기준으로 다시 변이 추정을 하여 최종적인 변이 벡터를 구한다. 이러한 방법으로 최상위

레벨에서 가장 낮은 레벨에 도달할 때까지 변이 추정을 하게 된다.

다해상도 분해를 기반으로 하는 변이 추정 방법도 블록을 기반으로 하기 때문에 적당한 블록 크기의 설정이 중요하다. 블록의 크기가 커지면 변이 벡터가 부정확해지고, 반대로 블록의 크기가 작아지면 변이 추정의 신뢰도가 감소하게 된다. 다해상도 기반으로 하는 변이 추정 방식에 있어서 블록 크기를 정하는 방식은 가변블록 크기방식과 고정블록 크기 방식 두 가지로 나누어 볼 수 있다. 본 실험에서는 모든 해상도에서 4x4, 8x8 그리고 16x16크기의 동일한 블록크기를 사용하는 고정블록 크기 방식을 사용하였다.

2.3 가려진 영역 추출 및 가려진 영역을 고려한 중간 영상 생성

한 영상에 대한 변이 정보를 나타내는 변이 맵을 가지고 중간영상을 생성하는 과정에서는, 가려진 영역의 판별이 선행되어야한다[4~10]. 가려진 영역이란 한 쪽 영상에서는 존재하지만, 다른 쪽 영상에서는 존재하지 않는 부분을 나타내는 것으로서, 변이 정보가 존재하지 않는 부분을 말하며, 좌 영상에서 보이지만 우 영상에서 보이지 않는 경우를 좌측 가려진 영역, 반대의 경우를 우측 가려진 영역이라 구분한다. 일반적으로 가려진 영역 추출을 위해서는 기준 좌 영상에서 우 영상으로 정합점을 찾아 이루어진 변이맵과 반대 방향으로 정합점을 찾아 이루어진 변이 맵을 비교하여 일대일 대응이 되지 않는 점을 가려진 영역이라고 판별한다.

가려진 영역을 구별한 변이 맵이 완성되면, 기준이 되는 좌우 영상과 변이 맵을 가지고 중간영상을 생성한다.

가장 기본적인 중간영상 생성은 단방향 보간법으로, 기준이 되는 좌우 영상간의 거리를 1이라고 했을 때, 생성하고자 하는 중간 시점의 위치를 설정하여 변이 맵에 가중치를 두고, 그에 해당되는 기준 영상의 화소 값을 가지고 보간한다. 하지만 일반적으로 단방향 보간법이 가진 오류가 있으므로, 이를 해결하기 위해 기준이 되는 좌우 영상을 모두 이용하는 확장된 양방향 보간법이 이용된다[4~10].

이러한 방법으로 변이의 정합점을 가진 영상에 대한 보간을 한 후에는, 가려진 영역에 대한 처리과정이 필요하다. 많이 이용되는 선형 양방향 보간법에서는 가려진 영역의 양 경계점이 가진 변이 정보를 가지고 좌우 가려진 영역을 판별하여, 좌측 가려진 영역일 경우에는 우 영상의 화소 값을, 우측 가려진 영역일 경우에는 좌 영상의 화소 값을 가지고 보간한다. 이 외의 방법으로, 가려진 영역을 보간할 때에 이러한 선형적인 보간법을 사용하지 않고 비선형적 방법을 사용하는 경우도 있다. 물체의 가려진 부분이 영상 평면에 대해 수평적이지 않을 경우 선형적 보간법에 따른 오류가 발생하거나, 이러한 물체가 겹치는 경우에 생기는 오류를 수정하는 방법들이 그 예이다[6,9].

다음은 구해진 완전한 변이 맵을 가지고 가려진 영역을 고려한 중간영상을 생성하는 방법을 서술한다. 일반적으로 중간영상 생성에는 어떤 방식의 보간법을 사용하는가 하는 것이 초점이 된다. 일반적으로 계산량과 효율성의 관계를 고려하여 선형 보간법을 주로 사용하며[5,7,10~12], 이 중

기준이 되는 좌, 우 영상이 모두 사용되는 양방향 보간법을 사용한다. 이 외에도 물체의 원형성을 고려한 비선형 보간법이나 객체 기반의 보간 방식에서 워핑등을 사용할 수 있으나 본 알고리즘에서는 가려진 영역을 고려한 선형 양방향 보간법을 사용한다.

중간영상의 선형 양방향 보간법을 설명한다. 중간영상을 얻기 위해서는 기준이 되는 좌우 영상과 이에 대응하는 변이 맵이 있어야 한다. 이 변이 맵은 앞에서 설명하였듯이 양방향 정합과 보정 과정을 통해 하나의 각 화소마다 하나의 대응점을 가지고 있다. 이러한 이유로 하나의 변이에 해당하는 중간영상의 화소 값은 단순히 기준 영상으로부터 해당하는 변이를 고려하여 보간한다. 다음 (7)은 변이의 정합점이 있는 중간 영상 시점의 보간법을 나타낸다.

$$I_n(i, j + \alpha * d_{L-R}(i, j)) = \alpha * I_L(i, j) + (1 - \alpha) * I_R(i, j + d_{L-R}(i, j)) \quad (7)$$

이때 α 는 좌우 영상의 거리를 1이라 했을 때, 생성하고자 하는 중간 시점이 좌 영상을 기준으로 얼마만큼 떨어져 있는 가를 나타내는 것으로서, 화 값을 할당할 때에 가중치의 역할을 한다. 이러한 가중치의 결과로 영상이 흐릿해지는 경우도 있으나, 대부분 오차를 줄이며 전체적으로 기준 영상간의 밝기에 차이가 있을 때에도 적절한 중간 시점 영상을 생성할 수 있는 결과를 나타낸다.

변이의 정합점이 있는 중간 시점 영상의 생성이 끝나면, 좌우 영상의 가려진 영역에 대한 보간법을 가지고 빈 공간을 채워 넣는다. 그림 6에서 빈공간 A,B점의 화소 값을 보간할 때에는, 점선으로 나타낸 α 값이 양쪽의 인접한 화소중에 이미 정합되어 보간된 화소의 변이를 비교하여, 스캔라인을 기준으로 이전의 L-R변이가 이후의 L-R변이보다 작은 것은 우영상에서, 그렇지 않은 것이 좌영상에서 화소의 값을 참조하여 보간한다.

이는 그림 6에서 보듯이 각 변이의 분포에 따라 좌우 영상의 가려진 영역을 판별할 수 있으며, 각각 대응하는 변이와 기준 영상을 가지고 보간하는 것이 적절하다는 이론에

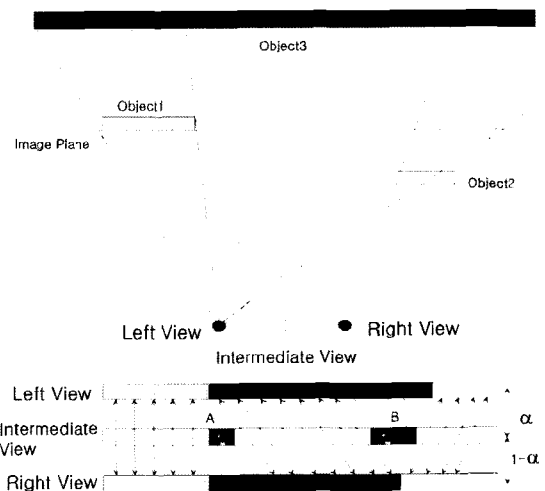


그림 6. 선형 양방향 보간 방식의 원리.
Fig. 6. The principle of linear bidirectional interpolation.

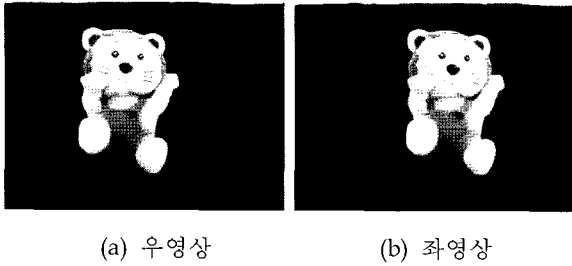


그림 7. 실험영상.
Fig. 7. The experimental images.

기인한다. 이와 같은 가려진 영역에 대한 보간법은 (8)에 나타나 있다.

$$I_{\alpha}(i, j + \alpha * d_{L-R}(i, j)) = \begin{cases} I_R(i, j + d_{L-R}(i, j)) & \text{if } d_{L-R_{best}} < d_{L-R_{\alpha}} \\ I_L(i, j) & \text{if } d_{L-R_{best}} \geq d_{L-R_{\alpha}} \end{cases} \quad (8)$$

VI. 실험 결과 및 고찰

실험은 펜티엄4 PC상에서 비주얼 C++를 이용하였다. 영상은 Sony XC-999 CCD카메라를 이용하여 베이스라인 거리를 3cm와 4cm 그리고 카메라와 물체와의 거리는 61cm로 하여 획득된 320×240크기의 영상을 사용하였다.

다해상도 기반 계층적 변이 맵을 이용하여 생성된 중간영상과 블록정합 알고리즘을 이용한 중간영상과의 비교를 하였으며, 기존의 깊이정보와 이미지의 기하학적 해석방법을 이용한 수직 축(Y축)을 기준으로 한 회전만을 고려한 중간영상 생성방법과의 비교 실험을 하였다.

최종적인 성능비교는 객관적 성능 비교와 주관적 성능비교를 사용하였다. 주관적인 성능비교는 실제 중간영상과 생성된 중간영상과의 차영상을 비교하였다. 그리고 객관적 성능비교는 PSNR을 비교하였으며, 사용된 PSNR(Peak Signal to Noise Ratio)은 (9)와 같이 나타낸다.

$$PSNR = 20 \log_{10} \left(\frac{255}{RMSE} \right) \quad (9)$$

여기서 RMSE(Root Mean Square Error)는 (10)과 같이 나타낸다.

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [R(i, j) - I(i, j)]^2} \quad (10)$$

(10)에서 R(i,j)는 실제 중간영상, I(i,j)는 생성된 중간영상을 나타내며, M,N은 영상에 포함하는 화소의 개수를 나타낸다.

1. 블록정합 단방향 변이맵을 이용한 중간영상생성
R-L 변이맵을 이용한 중간영상 생성결과를 보면 생성된 중간영상의 우측부분이 심하게 변형된 것을 볼 수 있었다. 이것은 중간영상 생성시 R-L 변이맵만을 이용하여 가려진 영역에 대한 고려가 없었기 때문에 생성된 중간영상에 많이 오류를 포함하고 있다.
2. 블록정합 양방향 변이맵을 이용한 중간영상 생성
그림 9와 그림 10은 가려진 영역을 고려한 양방향 블록

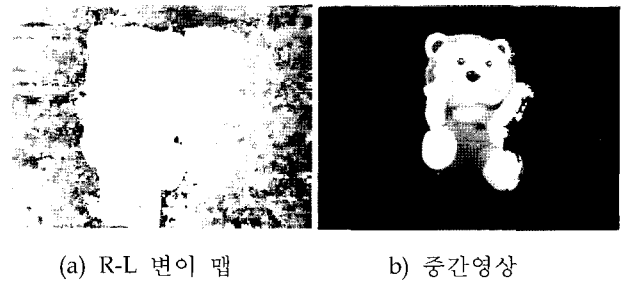


그림 8. 단방향 변이맵을 이용한 중간영상 생성.
Fig. 8. The intermediate image generation using one - directional disparity map.



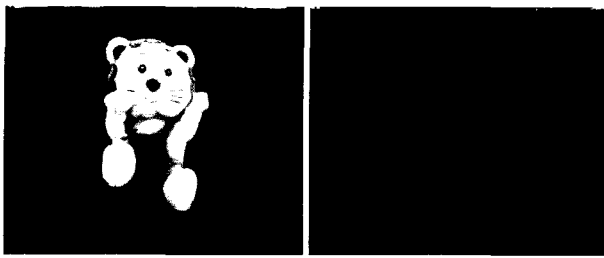
그림 9. 양방향 블록정합 변이맵을 이용한 중간영상.
Fig. 9. The intermediate image using bidirectional BMA disparity map (Baseline distance : 3cm).



그림 10. 양방향 블록정합 변이맵을 이용한 중간영상.
Fig. 10. The intermediate image using bidirectional BMA disparity map (Baseline distance : 4cm).

정합 변이 맵을 이용하여 생성된 중간영상이다. 중간영상 결과를 보면 단방향 변이 맵을 이용한 결과보다 생성된 중간영상의 오류가 적은 것을 볼 수 있다. 그러나 변이 정보가 정확하지 않기 때문에 생성된 중간영상에 오류가 발생함을 볼 수 있다.

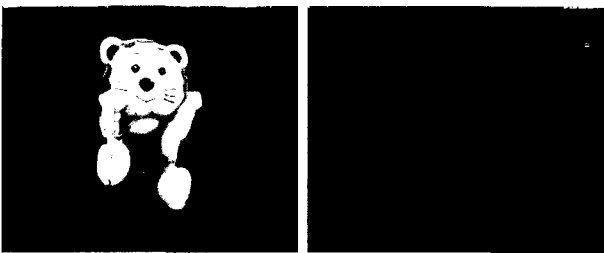
3. 다해상도 기반 계층적(HBMA) 변이맵을 이용한 중간영상 생성
그림 11과 그림 12는 다해상도 기반 계층적 변이 맵을 이용하여 생성된 중간영상이다. 블록정합 변이맵을 이용한 결과에 비해 차영상을 비교해보면 중간영상 생성결과에의 오류가 적게 나타났다.
4. 객관적인 성능평가 결과
표 1은 베이스라인 거리와 블록정합시 사용하는 블록의 크기에 따른 블록정합 알고리즘 변이맵과 다해상도 기반 계층적 변이맵을 이용한 중간영상 생성결과를 PSNR값으로 객관적인 성능평가를 수행한 결과이다. 표 1에서 블록정합 알고리즘 변이맵 보다 다해상도 기반 계층적 블록정합 변이



(a) 생성된 중간영상 (b) 차영상

그림 11. 양방향 계층적 변이맵을 이용한 중간영상.

Fig. 11. The intermediate image using bidirectional HBMA disparity map (Baseline distance : 3cm).



(a) 생성된 중간영상 (b) 차영상

그림 12. 양방향 계층적 변이맵을 이용한 중간영상.

Fig. 12. The intermediate image using bidirectional HBMA disparity map (Baseline distance : 4cm).

맵을 사용하였을 때 생성된 중간영상의 PSNR값이 0.1 ~ 1.4dB 더 좋은 값을 얻을 수 있었다. 그리고 중간영상 생성 시간도 다해상도 계층적 블록정합 변이맵을 사용하였을 때 7 ~ 10% 시간이 줄어들 수 있었다.

표 1. BMA와 HBMA 변이맵을 이용한 중간영상 생성 결과

Table 1. The result of intermediate image generation using BMA and HBMA disparity map

	Baseline	Block size	PSNR(dB)	TIME(s)
BMA	30mm	4 * 4	15.171	28
		8 * 8	18.382	101
		16 * 16	19.421	397
	40mm	4 * 4	14.422	29
		8 * 8	19.562	102
		16 * 16	22.840	400
HBMA	30mm	4 * 4	15.205	7.14% 감소
		8 * 8	19.850	6.93% 감소
		16 * 16	20.260	7.56% 감소
	40mm	4 * 4	14.292	10.34% 감소
		8 * 8	19.619	7.84% 감소
		16 * 16	22.667	6.75% 감소

5. 깊이정보를 이용한 중간영상 생성방법과의 비교

깊이정보를 이용하여 회전영상의 회전중간에서의 중간영상 생성방법[13]을 적용한 결과와 비교를 하였다. 본 논문에서는 평행식 스테레오 카메라 구조를 사용하였으나 기존의 방법[13]은 수직축(Y축)을 기준으로 회전된 영상에서의



(a) 생성된 중간영상 (b) 차영상

그림 13. 깊이정보를 이용한 중간영상 생성 결과.

Fig. 13. The intermediate image generation result using depth information of image.



(a) 생성된 중간영상 (b) 차영상

그림 14. 다해상도 기반 계층적 변이맵을 이용한 중간영상 생성 결과.

Fig. 14. The intermediate image generation result using HBMA disparity map.

중간영상 생성방법이다. 이 두가지 방법을 서로 비교하기는 어렵다. 그러나 기존방법에서 회전 각도가 아주 작다면 평행식 카메라 구조와 거의 차이가 없으므로 두 결과를 비교하였다. 기존의 방법 중 회전각도가 가장 적은 3° 영상의 중간영상 생성 결과와 본 논문에서 제안한 다해상도 기반 계층적 블록정합 알고리즘을 사용하여 생성된 중간영상 결과를 실제중간영상과의 PSNR값과 차영상으로 비교하였다.

깊이정보를 이용한 중간영상 생성결과의 PSNR값은 14.95dB 이고, 다해상도 기반 계층적 변이맵을 이용한 결과는 23.72dB의 결과를 얻었다. 깊이정보를 이용한 중간영상 생성방법에서는 깊이정보와 영상의 기하학적 해석을 통해 좌우 영상을 회전시킨 후 블록정합 알고리즘을 사용하여 두 영상을 정합시켜 중간영상을 만들었다. 그러나 이때 사용된 깊이정보가 정확하지 않기 때문에 PSNR 값이 낮게 나온 것으로 판단된다.

그림 13의 생성된 중간영상과 실제 중간영상과의 차영상을 보면 좌우 두 영상이 3°의 작은 회전각을 갖더라도 생성된 중간영상이 실제 중간영상에 비해 좌측으로 치우쳐있는 것을 볼 수 있으며, 생성된 중간영상도 많은 오류가 있음을 볼 수 있다. 그러나 그림 14의 다해상도 기반 블록정합 변이맵을 이용한 중간영상 생성결과는 영상의 오류가 적게 나타났다. 그리고 그림 13에서 배경부분의 깊이정보가 일정하기 때문에 생성된 중간영상의 배경이 깨끗하게 만들어 졌으나 그림 14의 경우 배경부분에서는 동일영역의 반복으로 정확한 변이추정이 어렵기 때문에 생성된 중간영상의 배경에 많은 오류를 포함하고 있다.

V. 결론

본 연구에서는 다해상도 기반 계층적 변이맵을 이용하여 중간영상을 생성하였다. 중간영상 생성시 양방향 변이맵을 이용하여 가려진 영역을 추출하였으며, 좌측 가려진 영역일 경우에는 우영상의 화소값 그리고 우측 가려진 영역일 경우에는 좌영상의 화소값으로 적절히 대체하여 중간영상을 생성하였다.

제안된 다해상도 기반 계층적 변이맵을 이용한 중간영상 생성방법이 기존의 블록정합 알고리즘을 변이맵을 사용하여 중간영상을 생성하는 방법에 비해 더 좋은 중간영상을 얻을 수 있음을 실제영상과의 차영상을 통해 주관적으로 확인할 수 있었고, 객관적 비교값으로 0.1~1.4dB정도 더 높은 PSNR값과 중간영상 생성시간이 6~10%이상 줄어들을 실험을 통해 확인하였다. 제안한 다해상도 기반 계층적 변이맵을 이용하고 가려진 영역을 고려한 중간영상 합성 방법은 양방향 변이맵을 이용하여 변이의 불연속점을 판별하는 과정을 통해 정확히 가려진 영역을 추출할 수 있으며 변이가 존재하는 정상적인 영역에서는 선형적인 보간법을 통해 보간하고, 가려진 영역에서는 좌, 우 영상으로부터 추정하여 보다 자연스러운 중간영상을 합성하였다.

향후 연구로는 세밀하고 정확한 변이맵을 생성하기 위해서 영상의 블록 경계 특성을 얻어내고, 다해상도 기법을 기반으로 변이추정시 블록의 경계특성에 따라 블록의 크기를 가변적으로 적용하는 방법의 연구가 필요하다.

참고문헌

[1] J. E. W. mayhew and J. D. Fisby, "Psychophysical and computational studies toward depth perception", *Artificial Intelligence*, vol 17, pp. 349-385, 1981.
 [2] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition : The Wavelet Representation," *IEEE Trans. Pattern Anal. and Machine Intell.*, pp. 674-693, July 1989.
 [3] S. ranganaath, "Image filtering using Multiresolution Representations," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 15, no. 12, pp. 1233-1240, Dec. 1993.

[4] J. S. McVeigh, "Efficient compression of arbitrary multi-view video signals," Ph. D thesis, Carnegie Mellon Univ., June 1996.
 [5] R. Lengagne, P. Fue. & O. Monga, "3D Stereo Reconstruction of human faces driven by differential constraints," *Image and Vision Computing* 18, pp. 337-343, 2000.
 [6] N. L. Chang & A. Zakhor, "A Multivalued Representation for View Synthesis," ICIP'99 Kobe, Japan, pp. 25-28, Oct. 1999.
 [7] J. S. McVeigh, M. W. Siegel & A. G. Jorddan, "Intermediate view synthesis considering occluded and ambiguously referenced image regions," *Signal Processing Image Communication*, pp. 21-28, Sep. 1996.
 [8] M. Maruyama & S. Kurumi, "Bidirectional Optimization for reconstructing 3D Shape from an Image sequence with Missing data," *Proceedings of the 1999 International conference on Image Processing*, vol. 3, pp. 120-124, 1999.
 [9] M. Luck & H. Schroder, "Error tolerant Interpolation of Intermediate Views for Real-time Applications," *Proceedings of the Stereoscopic Displays and Virtual Reality Systems VI*, pp. 208-218, Jan. 1999.
 [10] N. L. Chang & A. Zakhor, "View Generation for Three Dimensional Scenes from Video Sequences," *IEEE Trans. on Image Processing*, vol. 6, no. 4, Apr. 1997.
 [11] 양광원, "블록기반매칭을 이용한 깊이정보를 가진 영상의 중간영상생성," 단국대학교 석사학위 논문, 2001.
 [12] S. C. Joung, I. K. Hwang, S. J. Kim & J. K. Park, "Efficient 3D interpolation method for 3D reconstruction," *Proc. SPIE Three-Dimensional image Capture and Applications IV*, San Jose, January 2001.
 [13] 양광원, 허경무, 김장기, "깊이정보와 블록기반매칭을 이용한 스테레오 영상의 중간영상 생성," 제어·자동화·시스템공학 논문지, pp. 874-880, 2002.



허 경 무

1956년 9월 1일생. 1979년 서울대학교 전자공학과 졸업, 한국과학기술원 전기 및 전자공학과 석사(1981), 동대학원 박사(1989). 현재 단국대학교 전자컴퓨터학부 전자공학전공 부교수. 주요 연구분야는 3차원 디스플레이, 로

봇트 제어, 검사 자동화, 시스템 제어, 학습 제어, Servo Motor 제어, 디지털제어기 설계 및 시뮬레이션.



유 재 민

1973년 7월 15일생. 1999년 단국대 전자공학과 졸업. 동대학원석사(2003). 관심 분야는 머신비전, 영상처리.