# 변위 범위가 큰 다시점 비디오에 적합한 영상보간법

이천 오관정 호요성

광주과학기술원 정보기전공학부

{leecheon, kjoh81, hoyo}@gist.ac.kr

# A View Interpolation Method for Multi-view Video of Large Disparity

Cheon Lee Kwan-Jung Oh Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)

## 요약

차세대 방송서비스 개발의 일환으로 관심을 모으고 있는 다시점 비디오 부호화(multi-view video coding, MVC) 방식은 인접한 여러 대의 카메라로 동시에 획득한 영상을 효과적으로 압축하는데 그 목적이 있다. 이때, 중간시점 영상을 생성하여 부호화하는 과정의 참조영상으로 이용할 수 있으며, 이를 위해서는 다시점 비디오 특성에 맞는 영상보간 방법이 필요하다. 기존에 제안되었던 영상보간법은 변위의 검색범위를 초기에 설정하여 블록정합을 이용하여 화소 단위로 변위를 측정하기 때문에 카메라 사이의 거리가 크거나 객체의 움직임이 커서 변위의 변동이 심한 영상에서는 안정적인 화질의 영상을 얻기 어렵다. 또한, 고정된 크기의 블록을 이용하여 전체 변위를 측정하므로 객체의 변위차가 큰 영역에서 변위 오류가 많이 발생한다. 본 논문에서는 이와 같은 문제를 해결하여 보다 개선된 화질의 중간시점 영상을 얻기 위한 새로운 영상보간법을 제안한다. 영역분할을 이용한 초기의 변위측정 과정에서, 처음부터 최대 변위의 범위를 설정하는 대신에 블록 단위로 대략적인 변위를 측정한 후에, 가변 블록을 이용하여 보다 세밀한 변위를 측정한다. 이 방법은 변위차가 큰 객체의 경계 부분에서 보다 정확하게 변위를 측정할 수 있으므로, 화소 단위로 변위를 측정할 때 이전에 구한 변위 정보를 바탕으로 각 화소별로 검색 범위를 설정한다. 적응적으로 설정된 검색 범위를 이용하여 화소 단위의 변위를 측정하면 보다 개선된 변위를 얻을 수 있다. 추가적으로, 변위측정 과정에서 발생하는 변위의 오류를 최대한 줄이기 위해 각 단계별로 미디언 필터를 이용하여 변위 오류를 수정하였다. 본 논문에서 제안한 방법으로 실험한 결과 기존의 영상보간 방법보다 화질이 약 1~4 dB 정도 개선되었다.

## 1. 서론

최근 방송시스템의 기술이 발전함에 따라 디지털 기술의 특성을 이용한 방송 관련 부가 서비스들이 새롭게 선보이고 있다. 예를 들면, DMB(digital multimedia broadcast)와 같은 이동형 방송과 데이터 방송, 휴대 인터넷 방송 등이 상용화되었고, IPTV 및 T-Commerce 등의 다양한 방송 멀티미디어 서비스가 제공되기 시작했다. 또한, 디스플레이 장치의 기술이 발전하여 대화면 고화질의 방송서비스가 가능해졌다. 하지만, 이러한 기술은 TV 화면 자체가 2차원이기 때문에 화면상의 객체에 대한 거리감이나 질감을 제대로 느낄 수 없다.

3차원 비디오는 이러한 욕구를 충족시키기 위한 실감방송 서비스의 한 분야로서 여러 대의 카메라로 동시에 획득한 영상을 시청자가 현실감을 느낄 수 있는 영상으로 제작하고 입체적으로 디스플레이 하는 기술이다. 특히, 사용자가 자유로이 시점을 선택하여 시청할 수 있도록 다시점 비디오 데이터를 효과적으로 압축하는 다시점 비디오 부호화(multi-view video coding, MVC) 방법은 현재 MPEG 국제 표준화 기구를 중심으로 다각도로 연구되고 있다[1].

그 중에서 시점간 상관도를 이용하여 부호화 효율을 높일 수 있는 부호화 구조에 대한 연구, 시점간 조명 차이를 보상하는 방법, 부호화할 시점의 영상을 합성하고 이를 부호화에 적용하는 방법 등을 중점적으로 논의하고 있다.

인접한 시점의 영상들을 이용하여 중간 시점의 영상을 만들어내는 방법에 관한 연구는 계층적 깊이 영상(layered depth image)이나 라이트 필드 렌더링(light field rendering), 루미그래프(lumigraph) 등에서 이루어졌는데, 이들은 충분한 양의 입력데이터가 있어야 한다[2]. 하지만, 다시점 비디오는 여러 시점의 영상을 제공하는데 목적이 있기 때문에 이러한 기술들의 입력 데이터로는 충분하지 않다. 다시점 비디오에서 사용할 수 있는 데이터는 카메라 정보와 다시점 입력 영상인데, 이 데이터를 이용하여 중간시점 영상을 생성하기 위해서는 시점간의 변위를 이용하는 영상보간법이 적합하다.

기존에 다시점 비디오 부호화를 위하여 제안되었던 영상보간법은 블록 매칭을 이용하여 화소단위로 변위를 검색하는 방법이다. 그러나 이 방법은 초기에 변위 검색의 범위를 설정해야 하고, 변위차가 큰 영역에서 변위측정 오류가 많이 발생하는 한계가 있다. 본 논문에서는 이러한 한계를 극복하고 보다 나은 화질의 중간시점 영상을 얻기 위한 방법을 소개한다. 먼저 영역 분할을 이용한 초기 변위측정 기술을 통해

검색 범위를 전역부터 줄여가면서 변위를 측정한다. 또한, 가변 블록을 이용한 변위 측정과 적응적인 변위측정 범위를 갖는 화소단위 변위 측정을 통해 보다 정확한 변위를 측정하고, 각 단계별로 발생할 수 있는 변위 오류는 미디언 필터와 변위 정제를 이용하여 수정하였다.

## 2. 기존의 영상보간법의 한계

다시점 비디오 부호화에 적용하기 위해 지금까지 제안된 영상보간법으로 FTV의 구현을 목적으로 Droese가 제안한 방법이 있는데[3]. 이는 3차원 공간상의 한 점이 여러 대의 카메라로 투영될 때 그 빛의 세기는 같다는 점을 이용한다. 각 화소의 변위를 측정하고 그 변위를 이용하여 원하는 위치의 중간시점 화소값을 얻는다. 변위의 측정은 식 (1)과 같이 각 화소당 최소 비용값을 갖는 변위를 선택하는 WTA (winner takes all) 방법을 이용한다. 비용값은 식 (2)와 같이 변위의 변화에 따른 블록간 MAD(mean absolute difference)와 유사도 함수, 시간적 함수를 이용한다.

$$D(x,y) = \min_{d} C_W(x,j,d) \tag{1}$$

$$C_{W'}(x,y,d) = \overline{C_{W'}^{err}}(x,y,d) + \lambda \cdot C_{W'}^{reg}(x,y,d) + \gamma \cdot C^{temp}(x,y) \tag{2}$$

이때, $\lambda$는 유사도 함수의 가중치를 나타내는 상수이고, $\gamma$는 시간적 함수의 가중치를 나타내는 상수이다.

$$\overline{C_{W'}^{err}}(x,y,d) = \frac{\sum_{x,y \in W'} |I_L(x+\lfloor \alpha \cdot d \rfloor, y) - I_R(x+\lfloor (\alpha-1)\cdot d \rfloor, y)|}{|W|}$$

$|W|$ : 블록의 넓이 (3)

$$C_{W'}^{reg}(x,y) = \begin{pmatrix} D(x-1,y-1)+D(x,y-1) \\ D(x+1,y)+D(x-1,y) \end{pmatrix} / 4 \tag{4}$$

$$C^{temp}(x,y) = |D(x,y,t-1) - d| \tag{5}$$

식 (3)은 변위의 변이에 따른 블록간의 MAD를 구한 값이다. 유사도 함수는 식 (4)와 같이 주변에 있는 4개 화소의 변위를 이용한다. 비디오 데이터의 경우, 이전 화면의 변위와 현재 화면의 변위는 상관도가 높으므로 식 (5)와 같이 이용한다.

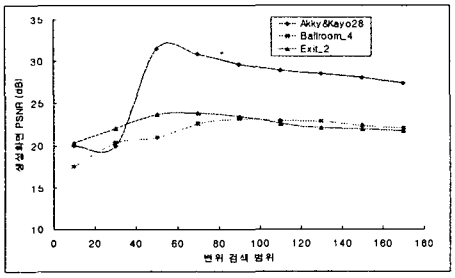


그림 1. 최대 변위 설정에 따른 화질 변화

위와 같은 방법은 기존의 스테레오 정합 기법과 같이 최대 변위를 설정하여 변위를 측정하므로, 최대 변위를 예측변위 측정에 정확히 적용하는 것이 중요하다. 그러나 비디오의 경우, 객체의 움직임을 예측하기 어렵고, 최대 변위가 자유로이 바뀔 수 있으므로, 최적의 최대 변위를 예측하는 것은 그리 쉽지 않다. 따라서 그림 1과 같이 변위의 검색 범위에 따라 화질이 크게 변하는 것을 알 수 있다. 또한, 기존의 방법은 화소 단위로만 변위를 측정하기 때문에 변위측정 오류가 많이 발생하여 변위 불연속점에서 오류가 전이되는 문제점을 가지고 있으며, 이 때문에 변위차가 큰 객체의 모양이 뒤틀리는 경우가 종종 발생한다. 이를 해결하기 위한 방법으로 가변블록 기반으로 변위를 측정하는 방법을 3장에서 제안한다.

## 3. 적응적인 변위검색 범위를 갖는 영상보간법

변위의 범위는 객체의 움직임과 다시점 카메라의 구조에 따라 유동적으로 바뀐다. 이러한 변화에 강인한 변위 측정을 위해 본 논문에서는 그림 2와 같은 방법을 제안한다.

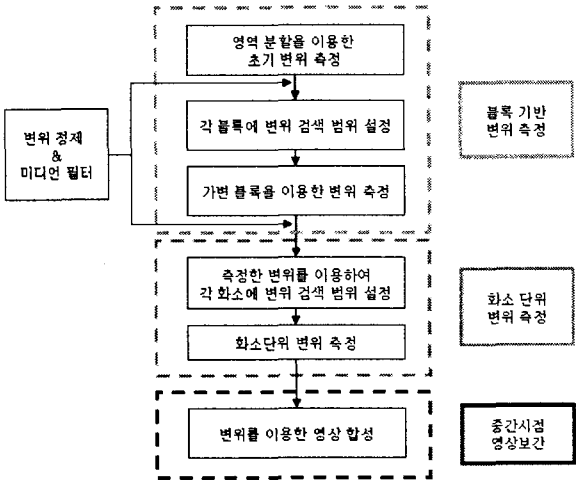


그림 2. 적응적인 변위 검색 범위를 갖는 영상보간법

### 가. 영역 분할을 이용한 초기 변위 측정

스테레오 영상은 객체의 위치를 제한하는 순서제한(ordering constraint)의 특징을 갖는데[4]. 이는 좌영상에서 객체 A가 객체 B의 왼쪽에 위치한다고 할 때, 우영상에서도 객체 A'는 객체 B'의 왼쪽에 위치한다는 조건이다. 이를 이용하면 좌우 영상을 특징점을 기준으로 영역 분할하고, 상응하는 영역 안에서 특징점이 강한 블록부터 변위를 검색할 수 있다. 본 논문에서 제안하는 방법에서는 특징점을 찾기 위해 좌우 영상의 그레디언트(gradient) 값을 이용하였다. 그레디언트를 이용하면 스테레오 영상의 조명 차이에 대한 영향을 적게 받고 특징점을 쉽게 추출할 수 있다.

### 나. 가변 블록을 이용한 변위 측정

블록 기반으로 변위를 측정할 경우에, 비슷한 변위를 갖는 역영의 변위는 비교적 정확하게 측정할 수 있으나, 객체의 경계와 같은 영역은 변위의 불연속점일 확률이 크므로 변위를 잘못 측정할 수 있다. 이러한 영역의 변위는 가변 블록을 이용하여 오류를 줄일 수 있다. 가장 큰 블록 단위로 변위를 측정한 후 동일한 변위를 이용하여 보다 작은 크기의 블록에 대한 비용값을 비교한다. 그림 3과 같은 과정으로 식 (6)을 이용하여 상위 블록의 비용값과 하위 블록의 비용값의 차이가 문턱치보다 클 경우는 하위 블록의 변위를 다시 측정한다. 이 방법을 이용하

면 변위차가 큰 경계부분에서 보다 정확한 변위를 측정할 수 있다.

$$\left|C_{W_{upper}}(i,j,d)-C^{i}_{W_{lower}}(i,j,d)\right|<T \tag{6}$$

이때, $W_{upper}=\sum_{i=0}^{3}W^{i}_{lower}$ 이다.

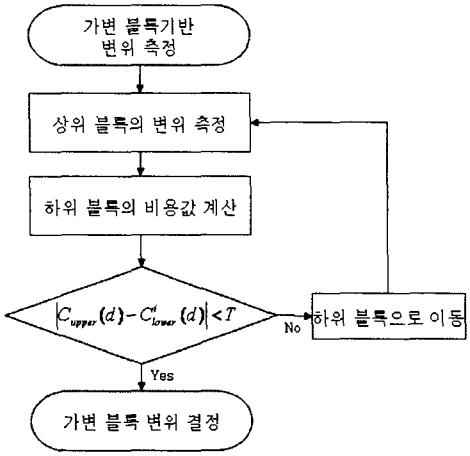


그림 3. 가변 블록을 이용한 변위 측정 과정

## 다. 화소단위 변위 측정

### (1) 화소단위의 변위 측정 범위 설정

블록 단위로 측정한 변위는 변위 정제와 미디언 필터를 이용하여 변위 오류를 줄일 수 있지만, 블록단위로 검색한 변위를 블록의 크기 때문에 화소간 변위 차이가 무시된다. 따라서 보다 정확한 화소 단위의 변위를 검색하기 위해 블록 크기보다 작은 범위를 이용하여 화소단위로 변위를 측정한다. 각 화소당 검색 범위는 식 (7)과 같이 이미 구해진 변위를 이용하여 재설정한다.

$$\begin{cases} MinRange = D(x,y)-Blocksize/2 \\ MaxRange = D(x,y)+Blocksize/2 \end{cases} \tag{7}$$

이때, $D(x,y)$는 블록기반으로 측정한 변위 지도이며, 새로 적용할 측정 범위는 블록 크기보다 작게 설정한다.

### (2) 화소단위의 변위 측정

앞 절에서 설정된 변위 측정 범위를 이용하여 화소 단위의 변위를 측정한다. 화소 단위의 변위 측정은 식 (8)과 같이 측정하고자 하는 위치의 점(x,y)을 둘러싼 블록의 MAD와 표준편차, 주변 화소의 변위를 이용한다. 이때 이용되는 비용값은 Veksler가 제안한 방법에 표준편차를 이용하는 방법을 식 (2)에 추가하여 만든 함수이다[5].

$$C_W(x,y,d)=C_W^{err}(x,y,d)+\lambda\cdot C_W^{reg}(x,y)+\rho\cdot C_W^{StDev}(x,y,d) \tag{8}$$

이때, $\lambda$는 유사도함수의 가중치를 나타내는 상수이고, $\rho$는 표준편차 함수의 가중치를 나타내는 상수이다.

## 라. 변위 오류 수정

정확한 변위를 측정하기 위해서는 각 단계별로 변위의 오류를 줄이는 것이 중요하다. 본 논문에서 제안하는 방법은 초기에 최대 변위를 설정하지 않고 영역분할을 이용하여 초기 변위를 측정하기 때문에 변위 오류가 많이 발생한다. 또한, 블록 기반으로 변위를 측정하면 변위가 비슷한 영역에서는 변위측정 오류의 확률이 감소하지만, 가변블록을 이용할 경우 최소 블록(2x2) 검색에서 변위 오류가 발생할 수 있다. 이러한 변위측정 오류는 보다 큰 블록을 이용하는 변위 정제와 잡음 제거에 유용한 미디안 필터를 이용하여 수정한다.

### (1) 변위 정제(Disparity Refinement)

변위 정제는 그림 4와 같이 블록보다 큰 크기의 정제 윈도우를 씌워 내부의 변위를 조사하고, 그 변위를 현재의 블록의 크기보다 1.5~2배 큰 블록에 적용하여 비용값을 적용한다. 그중 최소 비용값을 갖는 변위를 현재 블록의 변위로 설정한다. 이 과정을 거치면 블록 단위로 변위 오류를 줄일 수 있다. 특히, 영역분할을 이용한 초기변위 측정 후 블록 단위로 발생하는 변위측정 오류를 크게 줄일 수 있다.

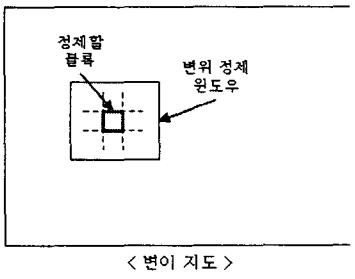


그림 4. 변위 정제

### (2) 미디언 필터(Median Filter) 이용

미디언 필터는 변위 지도에서 잡음과 같은 변위를 제거하는데 이용된다. 또한, 찾는 변위를 가능한 주변 변위와 유사하게 선정하므로 변위가 비슷한 영상에 대해 변위 오류를 크게 줄일 수 있지만, 변위차가 많은 영역에서는 오히려 주변의 변위에 오류를 증가시킬 수 있다. 이를 방지하기 위해, 화소 단위로 검색할 때 오류가 최대한 적게 발생하도록 미디언 필터의 윈도우 크기를 조절하는 것이 중요하다.

## 4. 측정한 변위를 이용한 영상합성

앞에서 측정한 변위를 이용하면 임의의 중간 시점에서 새로운 영상을 생성할 수 있다. 이때, 객체의 움직임과 조명의 차이가 선형적이라고 가정하면, 중간시점의 위치가 α일 경우에 각 화소의 값은 식 (9)와 같이 결정한다.

$$I_\alpha(x,y)=(1-\alpha)\cdot\hat{I}_L\left(x+\alpha\cdot d(x,y),y\right)+\alpha\cdot\hat{I}_R\left(x+(\alpha-1)\cdot d(x,y),y\right) \tag{9}$$

이때, $\alpha\cdot d(x,y)$와 $(\alpha-1)\cdot d(x,y)$가 정수가 아닐 경우, 인접한 화소의 값을 선형적으로 보간하여 $\hat{I}$를 각각 구하고, 그 값을 식 (9)를 이용하여 영상을 보간한다.

## 5. 실험 결과 및 분석

지금까지 제안한 방법을 카메라 사이의 거리가 20cm만큼 떨어진 'Exit', 'Ballroom', 'Akko&Kayo' 다시점 영상에 이용하였다. 이들 영상은 객체의 움직임이 활발하여 변위의 변동이 크다. 실험에서 가변 블록의 문턱값은 10~12로 설정했고, 식 (8)의 상수값은 $\lambda$=0.1, $\gamma$=3.1로 각각 설정하였다. 또한, 기본 블록의 크기는 16x16으로 설정하고, 가변 블록을 이용한 변위 측정시 2x2 블록까지 블록의 크기를 변화시켰다. 또한, 미디언 필터는 10x10의 윈도우를 이용하였다.

표 1은 변위 검색범위 설정에 따른 화질을 비교한 것이다. 실험영상을 이용하여 30 프레임의 합성영상을 얻고 그 화질을 원본 영상과 비교한 값이다. 제안한 방법의 검색 범위는 화소 단위로 변위를 측정할 때 적응적으로 검색 범위를 설정할 때 이용된다. 표 1의 결과를 비교해 보면, 기존의 방법보다 1~4dB 정도 화질이 개선되었음을 알 수 있다.

표 1. 변위 검색범위 설정에 따른 합성영상 화질 비교

| 테스트 영상 | 기존의 방법 | | | 제안한 방법 | | |
|---|---|---|---|---|---|---|
| | 최대 검색 범위 | | | 화소단위 검색 범위 | | |
| | 30 | 40 | 50 | 5 | 10 | 15 |
| Akko&Kayo | 27.778 | 31.453 | 30.959 | 33.013 | 32.659 | 32.273 |
| Ballroom | 20.699 | 21.043 | 21.435 | 25.287 | 25.297 | 25.265 |
| Exit | 21.350 | 22.270 | 23.205 | 24.525 | 23.675 | 23.672 |

단위: dB

그림 5는 'Akko&Kayo' 영상에 대한 각 화면별 합성영상의 화질을 비교한 것으로, 제안한 방법이 기존의 방법보다 전반적으로 1~2 dB 정도 개선되었음을 알 수 있다. 'Ballroom'이나 'Exit'의 경우에도 전반적으로 향상된 성능을 보였다. 특히, 제안한 방법은 블록 단위로 이미 구해진 변위를 바탕으로 변위를 측정하므로 최대 변위와 상관없이 작은 검색 범위를 이용한다.

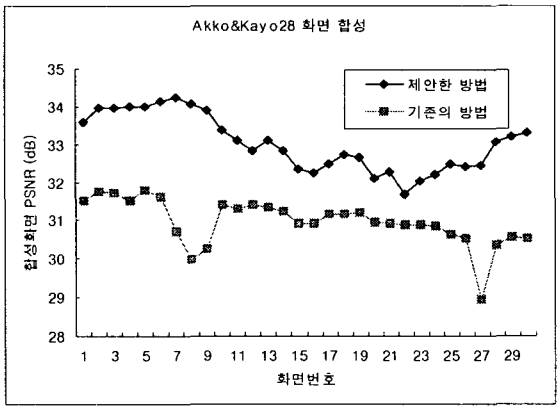


그림 5. 30프레임 합성 영상의 화질 비교

그림 6은 'Akko&Kayo28'영상의 첫 번째 화면의 합성 결과를 보인 것으로, 왼쪽에 위치한 사람의 얼굴의 경계 부분과 풍선의 경계 부분에서 왜곡이 발생했음을 알 수 있다. 이는 변위의 불연속점에서 변위 오류가 발생한 것이다. 표 1의 결과를 보면, 'Akko&Kayo'의 화질이 다른 두 영상보다 월등히 좋은 것을 알 수 있는데, 이는 영상의 특성상 변위의 종류가 적기 때문에 변위 오류가 적기 때문이다. 'Ballroom'이나 'Exit'는 변위의 종류가 많고 폐색 영역도 많기 때문에 화질이 비교적 낮은 결과를 보인다.

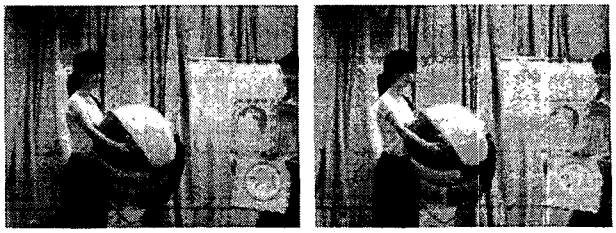


그림 6. 'Akko&Kayo28'의 원본영상과 합성영상

## 6. 결론

본 논문에서는 합성한 중간 시점의 영상을 다시점 비디오 부호화에 참조 화면으로 이용하기 위해 보다 개선된 화질의 중간시점 영상을 얻는 영상보간법을 제안하였다. 기존의 방법을 이용하여 합성된 화면이 최대 변위 설정의 영향을 크게 받는 반면에, 제안한 방법은 최대 변위를 설정하지 않고 보다 개선된 화질의 합성 영상을 얻을 수 있었다. 영역분할을 이용하여 대략적인 변위를 측정하고 가변블록을 이용하여 변위를 측정하여 변위의 불연속점에서 보다 정확한 변위를 측정할 수 있도록 했다. 최종적으로 이미 구한 변위를 기반으로 보다 정확한 화소 단위의 변위를 측정했다. 각 단계별로 변위를 측정할 때 발생하는 변위 오류는 변위 정제와 미디언 필터를 통해 수정하여 합성영상의 화질을 개선할 수 있었다. 실험결과에서 보듯이, 모든 영상에 대하여 제안한 방법이 보다 개선된 화질의 영상을 얻었다. 특히, 화소 단위로 변위를 검색할 때 변위측정 범위를 작게 해도 화질이 개선되었다. 본 논문에서 제안한 방법은 기존의 방법을 이용하여 얻은 영상의 최고 화질보다 약 1~4dB 정도 화질이 개선된 영상을 생성하였다.

## 감사의 글

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신부 ITRC 사업과 교육인적자원부 두뇌한국21(BK21) 정보기술사업의 지원에 의한 것입니다.

## 참고 문헌

[1] ISO/IEC JTC1/SC29/WG11 W8019, "Description of Core Experiments in MVC," April 2006.

[2] H. Y. Shum and S. B. Kang, "A Review of Image-based Rendering Techniques," *IEEE/SPIE VCIP*, pp. 2-13, June 2000.

[3] M. Droese, T. Fujii and M. Tanimoto, "Ray-Space Interpolation based on Filtering in Disparity Domain," *Proc. 3D Conference*, pp. 213-216, June 2004.

[4] A. L. Yuille and T. Poggio, "A Generalized Ordering Constraint for Stereo Correspondence," *MIT A.I. Laboratory Memo 777*, May 1984.

[5] O. Veksler, "Fast Variable Window for Stereo Correspondence using Integral Images," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 556 - 561, June 2003.