

영상 보간을 위한 유도 영상 필터링 기반의 변이 보정 기법의 성능 비교

신홍창, 이광순, 허남호
한국전자통신연구원

{ hcshin , gslee, namho } @etri.re.kr

A Comparison of guided image filtering algorithms based disparity
enhancement for view interpolation

Hong-Chang Shin, Gwang-Soon Lee, and Namho Hur
Electronics and Telecommunications Research Institute (ETRI)

요 약

본 논문에서는 영상 보간의 결과 측면에서 깊이를 보정하는 방법 중 하나인 유도 영상 필터링 기법을 비교한다. 실험을 위해 초기 깊이 영상을 두 종류의 유도 영상 필터링 기법으로 개선을 하였다. 초기 깊이 영상과, 각각의 필터링 기법에 의해 개선된 변이 영상을 이용하여 영상 보간을 하였고, 그 결과를 비교하였다. 결과로서 한 시점 영상의 텍스처 정보만을 이용하여 변이를 개선하는 유도 영상 필터링 기법으로 변이 영상을 개선하게 되는 경우에 육안으로는 구분이 갈 정도로 변이가 개선이 되지만, 영상 보간의 측면에서 보았을 때는 크게 차이가 없거나 오히려 품질이 저하되는 경우를 확인할 수 있었다.

1. 서론

무안경식 다시점 디스플레이는 두 시점 이상의 영상을 디스플레이 패널에 동시에 재현함으로써 안경을 쓰지 않아도 장면을 입체적으로 보다 현장감 있게 볼 수 있게 한다. 다시점 디스플레이에 필요한 시점의 수는 적게는 9 시점부터 많게는 수십 시점까지 디스플레이의 패널의 해상도와 디스플레이 방식에 따라 달라진다. 이렇게 많은 수의 시점의 영상을 생성하는 것은 CG(computer graphics)기반의 영상에서는 다시점 영상 렌더링을 통해 고품질의 영상을 생성하는 것이 가능하지만, 실사 기반의 영상을 생성하기 위해서는 수십 대의 카메라가 연결된 시스템과 카메라 간의 동기화 문제, 방대한 데이터 처리 등 선결되어야 할 문제들이 많다.

위와 같은 현실적인 문제로 인하여 한정된 시점 영상을 이용하여 인근의 가상 시점에서 영상을 생성하여 보다 많은 수의 시점을 생성하는 2D/3D 변환 혹은 스테레오/다시점 영상 변환 시스템과 같은 영상 합성을 이용하여 고품질의 다시점 콘텐츠를 생성하는 기술들에 대한 수요가 증대되고 있다.

스테레오 영상과 같은 한정된 시점을 이용하여 다시점 영상을 자동 생성하기 위해서는 크게 두 가지 문제가 해결되어야 한다. 첫 번째로, 주어진 시점에서의 3 차원 정보가 추출되어야 하며, 두 번째로는 추출된 3 차원 정보를 이용하여 DIBR(depth image based rendering)과 같은 영상 합성 기법을 통해 가상 시점에서의 영상을 정확하게 생성할 수 있어야 한다. 뿐만 아니라 생방송과 같은 실시간 응용사례를 고려한다면 위의 기술들이 실시간으로 구동이 되도록 고속화 구현 기법도 함께 연구가 되어야 한다.

3 차원 정보 추출은 주어진 원본 영상에서 장면 속의

객체와 배경의 깊이(depth) 정도를 추정하는 것으로서, 일반적으로 영상의 모든 화소에 대해서 두 영상 간의 서로 대응하는 대응점(correspondences)들을 찾는 스테레오 정합 기법을 주로 이용된다. 스테레오 정합은 십수 년동안 꾸준히 연구가 되어온 분야이며, 다양한 방법들이 제시되어 왔지만 다양한 환경에서의 강인성(robustness)과 실시간에 준하는 처리 속도를 모두 일정 수준 이상으로 만족시키는 방법은 여전히 난제로 남아있다.

깊이 영상(혹은 변이 영상)이 주어지면 장면의 기하관계를 토대로 영상 워핑(warping)을 통해 가상 시점에서의 영상을 생성할 수 있다. DIBR 기법이 널리 쓰이고 있는데, 각 시점 영상을 각각의 깊이 정보를 이용하여 가상 시점으로 화소를 투영시킨 후에 이를 다시 혼합(blending)하는 과정을 거친다. 이때, 원본 시점에서 보이지 않았던 영역이 홀(hole) 영역으로 드러나게 되는데, 이는 주변 정보를 이용하여 보간하는 인페인팅(inpainting) 과정을 통해 채워지게 된다. DIBR 과 같은 깊이 정보를 이용한 영상 합성 기법은 결과 영상의 품질이 깊이 영상에 크게 좌우된다.

최근 스테레오 정합의 연구 동향을 살펴보면, 전역(global) 알고리즘 방식에 비해 비교적 정확도가 낮았던 국부(local) 정합 알고리즘 방식이 다시 각광을 받고 있다. 국부 정합 방식은 영상을 블록 기반으로 나누어 블록 매칭(block matching)과 같은 국부의 비용을 계산하고, 후보 대응점들의 비용값을 근거로 가장 유사한 점 하나를 택하는 Winner-Takes-It-All 방식을 통해 대응점을 찾아낸다.

국부 정합 방식은 다음과 같이 크게 네 단계로 나뉘게 된다. i) 정합 비용 계산, ii) 비용 집적, iii) 변이 최적화, iv) 변이 개선. 국부 스테레오 정합 관련 많은 연구 사례들은 그동안 첫 번째부터 세 번째까지의 단계에 초점을 맞추고

있었으나, 최근에는 변이 후처리 개선 과정에 대해 활발하게 연구가 진행되어 국부 방식을 통해 초기 정합을 수행하고, 경계 보존 필터링을 통해 변이를 개선하는 방법들이 많이 제안되고 있다. Z. Ma 는 국부 스테레오 정합을 위한 가중치 중앙값 필터링 방법을 제안하였다[1]. 간단한 블록 기반 비용 직접 방법으로 추출한 초기 결과값을 가중치 중앙값 필터링으로 후처리 개선함으로써 계산복잡도가 높고 정확도가 높은 이용 집적 방식을 사용하는 다 스테레오 정합 기법 못지 않은 결과를 보여주었다.

Q. Yang 은 양방향 필터링(bilateral filtering)을 계층 구조를 사용하여 적응적 업샘플링(adaptive upsampling)을 하면서 경계부분을 보존하면서 변이 평활화 과정을 거치며 변이결과를 개선하는 방법을 제안하였다[2]. 이 방법은 계층적 구조를 사용하여 또한, 병렬처리가 용이하도록 설계가 되었기 때문에 하드웨어 기반으로 고속화가 용이한 장점이 있다.

위의 두 방법 모두 변이 영상을 개선할 때에 원본 영상을 유도 영상(guided image)으로 놓고 원본 영상에서 경계 부분을 추출하여 가중치를 계산하고, 구해진 가중치를 기반으로 경계 부분과 아닌 부분을 구분하여 평활화를 적용함으로써 변이 영상을 개선한다. 차이점은 전자의 경우 중심 화소의 인근의 변이값 중에 하나를 택일하는 중앙값 필터링을 하지만, 후자의 경우는 평활화 과정을 통해 변이값이 바뀌어 버리게 된다.

본 논문에서는 두 방식의 유도 영상 기반의 필터링(guided image filtering)을 이용하여 다시점 영상을 생성함으로써 각각의 방식에 따른 실험 결과를 보인다.

2. 시스템 구성도

이번 장에서는 스테레오 영상 기반의 다시점 영상 보간 기법에 대한 전체 과정에 대하여 설명한다. 그림 1 은 스테레오/다시점 변환의 전체과정을 보여준다.

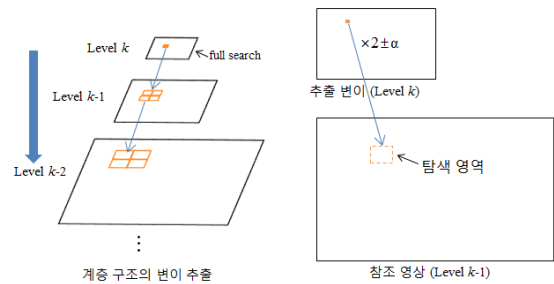


[그림 1] 스테레오/다시점 변환 개략도

스테레오 영상이 주어지면, 비교적 단순한 국부 스테레오 정합 기법을 이용하여 초기 스테레오 정합 결과를 얻는다. 여기서 스테레오 정합 단계의 목표는 실시간 추출을 위해서 시간복잡도와 품질 사이의 상충 관계(trade-off)를 고려하여야 하므로 완벽한 변이 영상을 추출하기보다는 영상 보간에 적합할 정도의 일정 수준 이상의 강인한 결과(robustness)를 보이는 변이 영상을 추출하는 것이다.

본 논문에서 택한 초기 스테레오 정합 방법은 계층 구조를 이용하여 국부 정합 기반으로 변이를 추출하는 기법이다. 이 방법은 짧은 시간 내에 비교적 정교하고 강인한 변이를 추출할 수 있어서 초기 스테레오 정합에 용이하다. 초기 스테레오 정합을 통해 얻어진 초기 변이 영상은 이후에 유도 영상 기반 변이 필터링 방법을 통해 개선되며, 개선된 변이 영상은 영상 보간(view interpolation)에서 3 차원 정보로 사용되어, 중간 시점에서의 영상을 생성한다.

2.1 초기 변이 추출



[그림 2] 계층 구조를 이용한 변이 추출

초기 변이 추출을 위해 계층 구조 기반의 블록 기반 스테레오 정합 기법을 사용하였다. 그림 2 는 계층 구조 정합을 도식화한 것이다.

영상 해상도에 따라 다르게 영상 피라미드를 구성한 후에 해상도가 가장 작은 최상위 계층에서부터 블록 단위로 국부 정합 방식으로 스테레오 정합을 수행한다. 최상위 계층에서 추출된 변이 값은 하위 계층의 탐색 영역 설정에 사용된다. 최하위 계층에 도달할 때까지 반복하며 최하위 단계에서의 변이 영상을 생성한다[5, 6].

2.2 변이 개선

초기 변이 영상이 주어지면, 후처리 변이 필터링이 수행된다. 본 논문에서는 두 종류의 유도 영상(guided image) 기반 변이 필터링을 사용하여 실험하였다.

가. 가중치-중앙값 변이 필터링

가중치-중앙값 필터링은 유도 영상 기반 필터링 방식의 하나로서 원본 영상에서 객체의 가장자리와 모서리 등의 경계부분을 검출하고 이를 보존하면서, 중앙값 필터링을 통해 이상점(outlier)을 제거한다.

이 방식의 기본 아이디어는 초기 변이 영상을 변이 별로 나누어 3 차원으로 변이층(disparity layers)을 구성한 후에 각 변이층에 해당하는 변이 영상마다 유도 영상 필터링(guided image filtering)[2]을 적용한다. 각 변이층의 경계 부분은 보존되고 나머지 영역은 평활화(smoothing)되는데, 이를 각 화소별 가중치 히스토그램(weighted histogram)으로 구성하여 중앙값에 해당하는 변이를 최종 변이로 선택하게 된다[1].

나. 고효율 양방향 변이 필터링

이 방식은 기존의 양방향 필터링(bilateral filtering)을 계층 구조를 이용하여 보다 고효율로 평활화하는 방법으로서 병렬처리 혹은 하드웨어 기반에서 높은 효율을 보여준다[3].

우선 계층 구조를 위해 영상 피라미드(image pyramid)를 구성하는데, 이때, 유도 영상을 다운샘플링(downsampling)할 때에 경계 부분을 유지하기 위해서 가장 가까운 거리의 화소의 값을 그대로 사용한다. 최하위 계층에 도달하면, 중심화소와 인접화소 간의 화소값 차이를 토대로 가중치를 계산하여 가중치과 가중치-변이값을 계산하고, 상위의 유도 영상 간의 화소값 차이를 토대로 다시 계층 간 가중치를 계산하여 적응적 업샘플링(adaptive upsampling)을 수행하며 평활화 과정을 거친다.

3. 실험 결과

그림 3 은 각 변이 영상에 따른 영상 합성 결과를 보인다. (c)는 초기 변이영상 결과이고, (d)는 (c)를 이용하여 합성한 결과 영상이다. 시점 위치는 좌/우 영상의 가운데로 정하였다. (e)는 고효율 양방향 변이 필터링(HEBF)를 적용한 변이 영상 결과이고, (f)는 (e)를 이용하여 합성한 결과 영상이다. (g)는 가중치-중앙값 필터링(WMF)를 적용한 변이 영상 결과이고, (h)는 (g)를 이용하여 합성한 결과 영상이다. (i)는 각 합성 영상의 부분을 비교한 영상이며, (i) 좌측 상하단 영상은 (d) 영상의 일부이며, 우상단 위치 영상은 (f), 우하단 영상은 (h) 영상의 일부이다.

본 실험에서 실험을 위해 사용한 중간 영상 합성 기법은 블록 기반의 가중치-평균 영상 합성 기법을 사용하였다. [5, 6]



(c) 초기 변이영상(좌)



(d) 합성 영상



(e) HEBF 후처리 변이 영상(좌)



(f) 합성 영상



(a) 좌영상



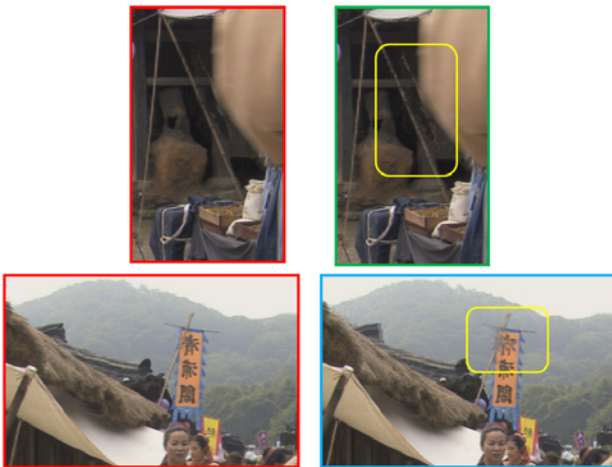
(b) 우영상



(g) WMF 후처리 변이 영상(좌)



(h) 변이 사용한 합성 영상



(i) 부분 화질 비교

[그림 3] 변이 영상에 다른 합성 영상 결과

표 1. 변이 추출 속도 비교[s]

알고리즘	소요 시간(s)
초기 변이 영상	0.016
HEBF 후처리 변이 영상	0.023
WMF 후처리 변이 영상(1/4)	0.150

<원본영상 해상도: 1280x720, Geforce GTX 690>

눈에 띄게 개선이 된 것을 확인을 할 수 있지만, 합성 결과는 크게 차이가 나지 않는 걸 확인할 수 있다. 오히려 일부 영역에서는 초기 영상을 사용한 경우가 더 나은 부분도 찾아낼 수 있다.

그 원인을 분석해보면, 초기 개선 영상의 변이는 일부 잡음이 포함될 수 있지만, 일련의 타당한 가정 하에 수행된 비용 집적 방식으로 추출된 대응점 간의 변이값이라고 볼 수 있는데, 유도 영상 기반의 변이 개선을 하게 되면, 한 시점의 원본 영상의 텍스처 성분을 이용하여 변이값을 변경하게 되므로 초기 변이값의 좌/우 일관성이 어긋나게 된다. 실험으로서 일관성 불일치의 수가 증가함을 확인하였고, 이는 원본 영상을 혼합(blending)하는 방식의 영상 합성 시에는 오히려 잡음 발생의 원인이 되었다.

4. 결론

본 논문에서는 각 변이추출 기법에 따라 추출된 변이 영상을 영상 합성 결과 측면에서 비교를 해보았으며, 유도 영상 기반의 변이 후처리 개선 기법을 사용한 경우에 변이 자체의 결과는 개선이 될 수 있지만, 영상 합성 결과 측면에서는 좌/우 일관성이 어긋나게 되어 잡음이 더 발생될 수도 있다는 실험을 통해 확인할 수 있었다.

감사의 글

본 연구는 2015 년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No. B0101-15-295, 초고품질 콘텐츠 지원 UHD 실감방송/디지털시네마/사이버니지 융합서비스 기술 개발)

참고문헌

- [1] Z. Ma, K. He, Y. Wei, J. Sun, and E. Wu, "Constant time weighted median filtering for stereo matching and beyond," in *Proc. of IEEE international Conference on Computer Vision (ICCV)*, 2013.
- [2] K. He, J. Sun, and X. Tang, "Guided Image Filtering," in *Proc. of European Conference on Computer Vision (ECCV)*, 2010.
- [3] Q. Yang, "High-Efficient Bilateral Filtering for Stereo Matching," in *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 36, no. 5, 2014.
- [4] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. of IEEE international Conference on Computer Vision (ICCV)*, 1998.
- [5] L. Zhang, D. Wang, and A. Vincent, "Adaptive reconstruction of intermediate views from stereoscopic images," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 1, pp. 102-113, 2006.
- [6] L. Zhang, "Fast stereo matching algorithm for intermediate view reconstruction of stereoscopic television images," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 10, pp. 1259-1270, 2006.

영상을 살펴보면 각 변이 추출 방법에 따라 변이 영상이