정규논문 (Regular Paper)

방송공학회논문지 제18권 제6호, 2013년 11월 (JBE Vol. 18, No. 6, November 2013)

<http://dx.doi.org/10.5909/JBE.2013.18.6.859>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

차폐영역 정보와 그리드 메쉬 기반의 영상 워핑을 이용한 다시점 영상 생성

임종명^{a)}, 엄기문^{b)}, 신흥창^{b)}, 이광순^{b)}, 허남호^{b)}, 유지상^{a)†}

Multi-view Image Generation using Grid-mesh based Image Domain Warping and Occlusion Region Information

Jong-myeong Lim^{a)}, Gi-mun Um^{b)}, Hong-chang Shin^{b)}, Gwangsoon Lee^{b)}, Namho Hur^{b)}, and Jisang Yoo^{a)†}

요약

본 논문에서는 스테레오 영상에서 얻은 다양한 특징들과 차폐영역에 대한 정보들을 활용하여 그리드 메쉬(grid-mesh) 기반의 영상 워핑 기법을 통해 다시점 영상을 생성하는 방법을 제안한다. 제안된 기법에서는 먼저 주어진 스테레오 영상에서 영상 특징 지도(image saliency map), 직선 성분(line segments) 그리고 변이 특징 지도(disparity saliency map)를 추출하고, 추출된 특징들에 대하여 품질을 향상시키는 과정을 거친다. 이 과정은 두 가지 단계로 나뉘는데, 먼저 차폐영역에 대한 정보를 활용하여 객체의 경계 부근에서 추출된 변이 특징 지도의 신뢰도를 향상시킨다. 다음으로 스테레오 영상에서의 시간적 일관성(temporal consistency)에 대한 정보를 활용하여 추출된 영상 특징들의 시간적 일관성을 높인다. 이렇게 품질이 향상된 특징 성분들을 활용하여 그리드 메쉬 기반의 영상 워핑 기법을 통해 다시점 영상을 생성한다. 실험 결과를 통해 제안된 기법으로 생성한 다시점 영상의 주관적 화질 측면에서 기존의 다시점 영상 생성 기법들보다 우수한 것을 확인할 수 있었다.

Abstract

In this paper, we propose an algorithm that generates multi-view images by grid-mesh based image domain warping using occlusion mask and various image features obtained from the stereoscopic images. In the proposed algorithm, we first extract image saliency map, line segments and disparity saliency map from stereo images and then get them through a process that improves the quality of extracted features. This process is accomplished in two steps. In the first step, reliability of disparity saliency map on object boundary regions is enhanced by using occlusion information. And in the second step, we enhance the quality of image features in terms of temporal consistency by using temporal consistency information for stereo images. With these enhanced features, multi-view images are generated by grid-mesh based image domain warping technique. Experimental results show that the proposed algorithm performs better than existing algorithms in terms of visual quality.

Keyword : View synthesis, Image Domain Warping, Multi-view generation, Occlusion information

a) 광운대학교 전자공학부(Kwangwoon University, Department of Electronic Engineering)

b) 한국전자통신연구원(Electronics and Telecommunications Research Institute, ETRI)

※ 본 연구는 방송통신위원회 및 한국방송통신전파진흥원의 ETRI연구개발지원사업의 일환으로 수행되었습니다. [11-921-02-001, 무안경 다시점 3D 지원 UHD TV 방송 기술 개발]

† Corresponding Author : 유지상(Jisang Yoo)

E-mail: jsyoo@kw.ac.kr

Tel: +82-02-940-5112

Manuscript received 14, June 2013 Revised 2, September 2013 Accepted 24, October 2013

1. 서론

2012년 지상파 아날로그 방송의 종료를 기점으로 디지털 방송의 시대가 열렸다. 네트워크 대역폭의 발전으로 인해 Full HD 급의 고화질 방송 시청이 대중화 되고 있고, IPTV 서비스와 스마트 TV의 보급으로 인하여 고품질의 다양한 콘텐츠를 필요로 하게 되었다. 또한 3D 입체 영상 및 자유 시점 영상 콘텐츠에 대한 관심도 발전하고 있다. 현재 우리나라를 포함한 다수의 국가에서 이미 3DTV 방송이 서비스 되고 있으며, 다양한 3D 입체 콘텐츠가 서비스되고 있다. 또한 노트북, 카메라, 프로젝터, 핸드폰 등 다양한 개인용 단말도 3D 콘텐츠를 이용할 수 있는 기능이 추가됨으로써 향후 3D 콘텐츠의 수요는 더욱 증가할 것으로 예상된다. 하지만 지금까지의 3D 서비스는 주로 양안 입체영상을 제공하는 형태가 대부분이며 따라서 안경을 착용해야 하는 불편함이 존재하고, 또한 하나의 시점 영상만 제공하기 때문에 사용자의 위치에 따라 입체감이 변하지 않는다는 단점이 존재한다. 이러한 단점을 보완하기 위해서는 자연에서 3D 입체 공간상의 물체를 볼 때 느끼는 완전 입체감을 느낄 수 있고 시점 변화가 자유로운 다시점 3D 서비스에 대한 기술개발이 필요하다. 다시점 무안경식 3D입체 서비스는 양안식 입체 영상과는 달리 다수의 시점으로 구성되어 있어 시야각이 넓다는 장점과 여러 방향에서 많은 사람이 동시에 시청하여도 뛰어난 입체감을 자연스럽게 느낄 수 있다는 장점이 있다. 2017년에는 이러한 다시점 3D 영상의 시험방송이 진행될 예정이며, 이에 따라서 관련 연구가 활발하게 진행되고 있다.

다시점 3D 입체 영상을 획득하는 방법은 크게 세 가지로 나눌 수 있다. 첫째, 필요한 시점 수만큼의 카메라를 사용하여 직접 다시점 영상을 획득하는 것이다. 이 경우 여러 대의 카메라 동기화 및 보정 작업이 쉽지 않다는 문제가 있다. 둘째, 컬러 카메라와 이에 대응되는 깊이 카메라(depth camera)를 사용하여 동시에 획득된 컬러 및 깊이 영상을 이용하여 DIBR(depth image based rendering)기법을 통해 다시점 영상을 생성하는 것이다^[1,2,3,4]. 마지막방법은 스테레오 카메라로부터 좌, 우의 영상을 획득하고 영상 워핑(image domain warping, IDW) 기법을 통해 다시점 영상을 생성하는 것이다^[5,6]. 현재 보급되고 있는 3D 디스플레이는 대부분 양

안식이기 때문에 대부분의 콘텐츠도 좌, 우 두 개의 영상으로 구성된 스테레오 영상이 대부분이다. 따라서 향후 무안경식 다시점 3D 디스플레이가 보급되면, 스테레오 영상을 입력으로 하여 DIBR 기법이나 IDW 기법을 통해 다시점 영상을 생성하는 기법이 매우 중요할 것이다^[2,3,4,5,6].

그러나 스테레오 정합 기법을 사용하여 정교한 변이지도(disparity map)나 깊이지도(depth map)를 추출하는 작업은 시간과 노력에 비하여 정확도가 많이 떨어지는 문제점이 있다. 따라서 차폐영역(occlusion)과 부정확한 변이 정보 인하여 생성된 가상시점 영상에 경계 잡음(boundary noise) 및 홀(hole)이 존재할 수도 있다. 이 경우는, 높은 성능의 홀 채움 및 경계 잡음 제거 기법이 필요하다^[7]. 그러나, 자동화된 홀 채움 및 경계 잡음 기법에서는 정교한 변이지도 혹은 깊이지도를 획득하는 것이 어렵기 때문에, 신뢰도를 보장할 수 없다^[8].

이러한 문제들에 대한 해결책 중 하나로 영상기반 워핑 기법을 활용하여 다시점 영상을 생성하는 기법이 제안되었다^[5]. 이 기법은 스테레오 영상에서 화소 전체에 대한 변이 정보를 추출하지 않고, 광류(optical flow)^[9]와 SIFT(scale-invariant feature transform)^[10]를 활용하여 특징점에 대한 변이만을 구한다. 또한, 스테레오 영상에서 직선 성분(line segment) 및 명암변화 특징(intensity gradient saliency)^[6,11]을 추출하여 가상 시점에서도 이러한 특징들이 유지되도록 하는 제약 조건으로 활용한다. 그러나 이 기법도 광류와 SIFT의 부정확한 변이 추정 값으로 인하여 객체 경계 영역에서의 특징점들의 변이가 낮은 신뢰도를 갖는 단점을 가지고 있다. 그리고 화소값의 미세한 변화로 인하여 영상의 특징들이 프레임별로 추출되는 특징이 달라질 경우, 해당 부분의 가상 시점 영상에서 흔들림 현상이 발생할 수 있다.

본 논문에서는 앞서 언급했던 IDW기법^[5]의 문제점들을 개선하기 위하여 차폐영역 추출 정보와 시간적 일관성 조사를 이용한 새로운 IDW 기법을 제안한다. 제안한 기법은 먼저 차폐영역에 대한 마스크 정보를 활용하여 객체 경계 영역에서의 낮은 신뢰도의 변이 값들을 신뢰도가 높은 변이 값들로 품질을 개선하는 과정을 거쳐 변이의 신뢰도를 높인다. 차폐영역에 대한 정보는 원본 좌(우)영상과 IDW를 기법을 통해 각각 생성된 가상의 좌(우)영상 간 합성 오차

계산^[12]을 통하여 이 정보를 획득한다. 다음으로 프레임간의 시간적 일관성(temporal consistency) 대한 정보를 이용하여, 스테레오 영상에서 추출된 특징 성분들이 시간적 일관성을 갖도록 보완한다.

본 논문의 구성은 다음과 같다. I장 서론에 이어, II장에서는 제안하는 기법의 전체적인 구조와 세부적인 과정을 설명한다. III장에서는 제안한 기법의 성능 비교 실험결과를 보여주고, 마지막으로 IV장에서 결론을 맺는다.

II. 제안하는 다시점 영상 생성 기법

이 장에서는 본 논문에서 제안하는 IDW(image domain warping) 기반의 다시점 영상 생성 기법에 대하여 설명한다. 그림 1은 본 논문에서 제안하는 IDW 기반의 다시점 영

그림 1. 제안된 다시점 생성 알고리즘의 흐름도
Fig. 1. Flow chart of the proposed multi-view synthesis algorithm

상 생성 기법의 흐름도이며, 세부적인 내용은 이어지는 하위 절에서 설명한다.

1. 스테레오 영상 내 특징 추출

제안하는 기법의 첫 번째 과정으로 스테레오 영상으로부터 변이 특징 지도(disparity saliency map)를 생성하는데, 이는 블록 단위의 광류(optical flow)^[9] 계산 기법과 SIFT(scale-invariant feature transform) 기법^[10]으로부터 계산한다. 광류 계산 기법은 다음의 세 가지 가정을 기반으로 하는 Lucas-Kanade 기법^[9]을 사용한다. 시간적으로 연속적인 동영상에서 동일한 객체의 밝기 값은 일정하게 유지된다는 밝기 항상성(brightness constancy)과 영상 내에서의 움직임은 그다지 빠르지 않다는 시간 지속성(temporal persistence), 그리고 공간적으로 서로 인접하는 점들은 동일한 객체에 속할 가능성이 높고 동일한 움직임을 갖는다는 공간 일관성(spatial coherence)이 그것이다. Lucas-Kanade 기법에서는 가우시안 피라미드(gaussian pyramid)를 적용하여 두 영상 간에 존재하는 화소의 움직임 정보를 구한다. 이 정보는 색과 모양 변화에 대하여 강인한 결과를 보여주며, 블록 정합 등의 기법에 비하여 정확도가 높다는 장점이 있다.

SIFT(scale-invariant feature transform) 기법은 물체의 크기나 방향에 상관없이 높은 정확도로 해당 물체의 위치를 추출할 수 있다. 이 SIFT 기법을 통해 좌, 우 스테레오 영상에 포함된 특징점을 찾고 각 영상에서 찾은 특징점들의 위치 차이를 이용하여 특징점들의 변이를 찾는다^[10]. 이렇게 구한 변이는 광류 계산을 통해 얻은 변이와 상호 보완적으로 사용한다.

그림 2는 SIFT 기법을 이용하여 추출된 스테레오 영상

그림 2. SIFT기법을 통해 추출된 특징점과 예측된 변이 정보

Fig. 2. Extracted Feature points and estimated disparity information by SIFT (a) left image (b) right image

내에서의 특징점 및 특징점 간 변이를 나타내고 있다. 녹색 선은 좌, 우 영상 간 특징점 정합을 통하여 영상 내에서 추출한 특징점의 변이를 나타내며, 녹색 선의 길이는 대응 특징점 간의 변이 크기를 의미한다.

지금까지 설명한 광류계산과 SIFT를 이용한 변이 추출 결과를 서로 결합하여 최종 변이 특징 지도를 구성한다. 이때, 동일한 특징점에 대해 추출된 변이가 두 기법 모두에 존재할 경우에는 신뢰도가 좀더 높다고 판단되는 SIFT기법에 의해 얻어진 변이를 사용하게 된다. 광류 계산 및 SIFT를 이용하여 이렇게 결합된 변이 특징 지도는 기법의 특성상 영상 내에 있는 객체의 경계영역에서 낮은 신뢰도를 갖는다. 본 논문에서는 입력받은 차폐영역(occlusion region) 추출 정보를 이용하여 이 영역에서의 변이 특징 지도 신뢰도를 높인다. 또한 본 논문에서 제안하는 기법에서는 다시 점 영상 생성 시의 영상 왜곡을 최소화하기 위해 명암변화 특징 지도(intensity gradient saliency map)^[11,12,13]와 직선 성분(line segment)^[14,15]도 추출하여, 영상 워핑 시 제한 조건으로 사용한다.

여기서 명암변화 특징 지도의 원리를 설명하면 다음과 같다. 인간의 인지과정은 시각체계를 통해 입력되는 영상 중 의미 있는 특징만 선택적으로 선별하여 인지함으로써 보다 빠르게 많은 처리를 수행할 수 있다. 이에 대한 연구는 생물학, 인지공학, 컴퓨터 비전 분야 등에서 활발히 진행되고 있다. 특히, 인간의 인지과정을 바탕으로 하는 특징 지도(saliency map)는 영상을 이진화하여 객체 영역과 객체가 아닌 영역으로 분리함으로써 관심 있는 주요 객체를 추출하는데 주로 사용되고 있다^[12]. 본 논문에서는 S. Montabone가 제안한 방법^[12]을 사용하여 명암 변화 특징 지도(intensity gradient saliency map: IGSM)를 구한다.

한편, 직선 성분은 영상 워핑 시 영상에 존재하는 직선 부분의 왜곡을 최소화하기 위한 목적으로 사용된다. 기존에는 Canny 에지 추출 기법과 Hough 변환이 많이 이용되었으나, 본 논문에서는 연산의 복잡도를 낮추기 위해 LSD(line segment detector) 기법^[15]을 사용하였다. 이렇게 추출된 명암 변화 특징 지도와 직선 성분은 그리드 단위의 매쉬

(a)

(b)

(c)

(d)

그림 3. 스테레오 영상으로부터 추출된 특징들 (a) 원본 좌 영상 (b) 변이 특징 지도 (c) 명암 변화 특징 지도 (d) 직선 성분

Fig. 3. Extracted features from stereo images (a) original left image (b) disparity saliency map (c) intensity saliency map (d) line segments

워핑(Grid-mesh warping) 기법으로 다시점 영상을 생성할 때 영상의 중요한 부분과 직선 부분에서 왜곡이 최소화 되도록 사용된다^[5,6].

그림 3(b), (c), (d)의 영상은 입력받은 'BeerGarden' 스테레오 영상으로부터 추출한 변이 특징 지도, 명암 변화 특징 지도, 직선 성분의 예를 각각 보여준다.

2. 차폐영역 추출 정보를 이용한 변이 특징 지도의 신뢰도 향상

앞 절에서 언급하였듯이 제안하는 기법에서 변이를 추출하기 위하여 사용하는 광류 계산 기법 및 SIFT 기법은 객체 경계영역에서 신뢰도가 낮다는 단점이 있다. 이는 가상 시점 생성 시에 객체의 왜곡을 발생시킬 수 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 입력받은 차폐영역 추출 정보를 활용하여 객체 경계영역에서의 변이 신뢰도를 향상시켜 가상 시점에서의 객체 왜곡을 최소화하도록 한다.

그림 4. 차폐영역 마스크의 예
Fig.4. Example of an occlusion mask

차폐영역 추출 정보는 그림 4와 같이 0(비차폐영역)과 1(차폐영역)로 구성된 마스크(mask)의 형태로 표현될 수 있다. 이 정보는 주어진 스테레오 영상에 대한 정교한 깊이지도가 있다면, LRC(left/right disparity consistency checking) 기법을 사용하여 구할 수 있다. 그렇지 않다면, 원본 좌(우) 영상과 IDW를 통해 생성된 가상의 좌(우)영상의 합성 오차(왜곡)가 많은 영역^[16]으로부터 얻는다.

(a)

(b)

(c)

(d)

그림 5. 'BeerGarden' 시퀀스의 변이 특징 지도 및 컬러 영상과의 중첩 영상 (a) 수정 전의 변이 특징 지도 (b) 수정 전의 중첩 영상 (c) 수정 후의 변이 특징 지도 (d) 수정 후의 중첩 영상

Fig. 5. Overlapped images of color image and its disparity (a) before modification (b) after modification

그림 5는 'BeerGarden' 시퀀스로부터 추출한 변이 특징 지도 및 컬러 영상과 중첩한 영상을 나타낸 것이다. 그림 5(a)와 (b)에서, 변이 특징 지도와 컬러 영상의 경계들이 잘 맞지 않음을 확인할 수 있다. 일반적으로 차폐영역은 주요 객체와 배경의 경계부분에서 발생하기 때문에, 제안하는 기법에서는 차폐영역 정보를 주요 객체의 경계영역에 대한 정보로 활용하여 추출된 변이 특징 지도의 객체 주변에서의 신뢰도를 높였다.

그림 6은 본 논문에서 변이 특징 지도의 신뢰도를 향상시키기 위해 사용한 방법이다. 그림 6에서, 녹색과 파랑색으로 표시된 영역은 객체 경계에서의 배경 영역 및 객체 영역을 나타낸 것으로 차폐영역 정보를 객체의 경계영역에 대한 정보로 활용하여 추출하였다. 제안하는 기법에서는 객체 경계 영역에서 변이 특징 지도의 신뢰도를 높이기 위해서, 이 영역들을 기준으로 일정 거리 내에 있는 변이를 수정하게 된다. 또한 연한 색 블록은 수정에 사용되는 변이를 나타낸 것이며, 진한 색 블록은 수정할 변이를 나타내는 것으로 연한 색 블록들의 평균값으로 수정된다. 이 과정은 그림 6에서 표시한 화살표 방향처럼 경계의 바깥 영역에서 경계 영역의 방향으로 진행된다.

그림 6. 추출된 변이의 신뢰도 향상
Fig. 6. Enhance reliability of extracted disparity

3. 추출된 특징들의 시간적 불연속성 정제

카메라를 통해 연속된 영상을 촬영할 때, 카메라가 고정되어 있어도 조명 변화 등의 이유로 움직임이 없는 배경 영역에서의 미세한 화소 컬러값의 변화가 발생할 수 있다.

제안하는 기법에서 특징 성분들을 추출할 때도 화소 컬러값의 미세한 변화로 움직임이 없는 영역에서 프레임 간 추출된 특징 성분들의 차이가 발생할 수 있다. 그림 7은 이러한 현상을 보여주는 예시로, 전체적으로 움직임이 거의 없는 'Samgyetang' 시퀀스의 1~3번 프레임에서 추출한 직선 성분들을 각 프레임마다 다른 색으로 표현한 후 중첩하여 나타낸 것이다.

그림 7. 연속된 프레임에서 추출된 직선 성분들의 중첩 영상(1~3프레임, '삼계탕' 시퀀스)

Fig. 7. Overlapped line segments extracted from successive frame(1~3 frames, 'Samgyetang')

그림 7에서 보는 바와 같이 직선 성분들이 여러 가지 색으로 표현된 것을 확인할 수 있는데, 이는 밝은 녹색으로 표현된 직선 성분들만이 세 프레임 모두에서 추출된 것이며, 그 외의 다른 직선 성분들은 하나 혹은 두 프레임에서만 추출된 것이다. 그림 8은 'Samgyetang' 시퀀스의 첫 번째와 두 번째 프레임의 차이 영상을 나타낸 것으로, 그림 7과 그림 8(a)를 보면 연속된 프레임 간의 움직임이 나타나지 않는 부분에서 추출된 직선 성분들이 차이를 보이는 것을 확인할 수 있다. 이러한 움직임이 없는 영역에서의 특징 성분 차이는 가상시점 생성 시에 수행하는 최적화 과정에서 다르게 작용하여 그림 8(b)와 같이 생성된 가상시점 영상의 프레임간 차이를 발생시키고, 결과적으로 화면이 흔들리는 현상이 나타날 수 있다. 따라서 영상 내의 움직임이 없는 영역에서 추출된 특징의 시간적 일관성을 보장하기 위한 후처리가 필요하다.

본 논문에서는 시간적 불연속성 정제를 위하여 블록 기

그림 8. 연속된 프레임간의 차 영상 (첫 번째와 두 번째 프레임, '삼계탕'시퀀스) (a) 원본 좌 영상 (b) 시간적 일관성 보상 전의 가상 시점 영상 (c) 시간적 일관성 보상 후의 가상 시점 영상

Fig. 8. Difference image of successive frame (1st frame and 2nd frame, 'Samgyetang' sequence) (a) original left image (b) virtual view image without temporal inconsistency refinement (c) virtual view image with temporal inconsistency refinement

반 MAD(mean absolute difference)를 활용하여 영상 내의 움직임을 검출 한다^[17]. 먼저 좌, 우 영상에 대하여 RGB 각 성분별로 블록 단위의 MAD를 계산하고, 계산된 성분별 MAD값이 모두 임계값보다 작을 때만 움직임이 없는 블록으로 판단한다. 움직임이 없는 영역에서는 이전 프레임에서 추출된 특징 성분을 그대로 사용하며, 그 외의 영역에서는 현재 프레임에서 추출된 특징 성분을 사용하여 최적화 과정을 진행한다. 시간적 불연속성을 정제한 특징 성분을 이용하여 생성한 가상 시점의 차 영상인 그림 8(c)를 보면 정제 전에 비하여 연속된 프레임간 컬러차이가 많이 줄었음을 확인할 수 있다.

4. 그리드 메쉬 워핑을 활용한 다시점 영상 생성

본 논문에서는 앞 절에서 추출된 특징들을 사용하여 그리드 메쉬 기반의 워핑 기법(grid-mesh based image domain warping technique)을 적용하여 다시점 영상을 생성하게 된다. 그리드 메쉬 기반의 워핑 기법은 그림 9와 같이 영상을 일정 크기의 메쉬(mesh)의 형태로 구성한 후, 각각의 메쉬에 해당하는 변이 값을 기반으로 하여 영상을 메쉬 단위의 워핑을 수행하여 전체적인 영상을 재구성하는 기법이다^[5, 6]. 이 때, 기반이 되는 변이 값은 추출 및 경계영역에서 신뢰도가 향상된 변이 특징지도로 사용한다.

그리드 메쉬 기반 워핑 시 앞 단계에서 추출된 특징들을 이용하여 영상 내의 고유한 특징들이 가상 시점에서도 최대한 유지되도록 각 메쉬들에 대하여 최적화 과정을 거친다. 메쉬의 최적화 과정은 영상의 외곽부분을 제외한 각 메

그림 9. 그리드 메쉬로 표현한 영상의 예시 ('Story' 시퀀스)

Fig. 9. The example of image expressed by grid-mesh ('Story'-sequence)

쉬들의 꼭지점에 적용되는데, 그 과정은 다음과 같다. 먼저, 변이 값들을 기반으로 원본 좌(우) 영상의 메쉬들을 변형시켜 가상의 우(좌) 영상을 만든다. 그 후, 변형된 메쉬의 꼭지점을 기준으로 하여 일정 범위 내에서 위치를 변화시키며 최적의 가상의 우(좌) 영상을 갖도록 하는 꼭지점의 위치를 지정한다. 이 때, 가상의 우(좌) 영상과 원본 우(좌) 영상의 식 (1)로 표현되는 에너지를 계산하여 최소의 에너지를 갖는 꼭지점의 위치로 선택하게 된다^[18,19,20].

$$E_w = \lambda_c E_c + \lambda_i E_i + \lambda_t E_t \quad (1)$$

E_c 는 가상의 영상과 원본 영상의 차이를 나타내고, $E_i \cdot E_t$ 은 가상의 영상과 원본 영상에서의 명암 변화 · 직선 성분들의 차이값을 나타내며, λ 는 각 에너지들의 가중치를 나타낸다. 따라서 최종적으로 합산된 에너지 E_w 가 작을수록

스테레오 영상으로부터 추출된 특징 성분들이 가상 시점으로 이동했을 때 얼마나 유지되었는가를 의미하게 된다.

이러한 최적화 과정을 거쳐서 원본 좌(우) 영상에서의 가상의 우(좌) 시점을 생성할 때, 각 메쉬들의 최적의 형태가 결정된다. 이렇게 결정된 최적의 메쉬를 통하여 가상 시점 영상을 생성한다. 좌, 우 영상 사이에 임의의 가상 시점 영상을 생성할 때, 생성할 가상 시점과 원본 영상 사이의 거리와 좌, 우 영상의 거리 비율만큼 최적화된 메쉬의 형태를 조정하여 가상 시점에서의 메쉬의 형태를 결정한다.

또한 가상 시점 영상 생성 시, 차폐영역에 대한 정보를 활용하여 좌-우 영상에서 나타나는 차폐영역을 고려한다. 차폐영역 정보와 변이 정보를 기반으로 식 (2)~(4)을 이용하여 각 가상 시점에 해당하는 차폐영역 추출 정보를 생성한 후, 영상기반 위핑(IDW) 기법을 적용하여 가상 시점 영상을 생성한다.

$$\alpha = \min(|D_{V-L}|, |D_{V-R}|) / |D_{L-R}| \quad (2)$$

$$L = \alpha L_0 \quad (3)$$

$$x = x_0 + \alpha d \quad (4)$$

여기서 α 는 가상 시점과 가까운 원본 시점과의 거리를 나타내는 것으로 D 는 각 시점 사이의 거리를 나타낸다. L 와 x 는 가상 시점에서의 차폐영역의 길이 및 위치를, L_0 와 x_0 는 입력받은 차폐영역의 길이 및 위치를 나타내며 d 는 해당 차폐영역 주변 객체의 변이를 나타낸다. 가상 시점에서 차폐영역을 제외한 영역은 원본 시점 영상 중 가상 시점

으로부터 가까운 시점의 영상으로부터 생성된다. 예를 들어 생성되는 가상 시점의 위치가 왼쪽에 가깝다면, 해당 영역은 좌 시점 영상으로부터의 위핑을 통해 생성된다. 반대로, 가상 시점에서의 차폐영역은 가상 시점으로부터 먼 시점의 영상으로부터의 위핑을 통해 생성된다.(이 경우 우 시점에 해당)

다음의 그림 10은 'BeerGarden' 스테레오 영상을 입력받아 제안된 기법을 이용하여 생성한 8개의 가상 시점 영상과 입력 좌측 영상 등 총 9개 시점의 영상을 보여준다.

III. 실험 결과

본 논문에서 제안한 다시점 영상 생성 기법의 성능을 평가하기 위해 MPEG에서 제공하는 다시점 영상 시퀀스인 'BeerGarden' (1,920x1080)과 'Cafe' (1,920x1080)를 테스트 영상으로 사용하였다. 그림 11(a)와 (b)는 각각 제안된 기법과 DIBR(depth image based rendering) 기법[4]에 의해 생성된 'Cafe' 영상의 가상 시점 영상으로, 좌 영상과 우 영상을 포함한 총 10개의 시점 중 다섯 번째 시점 영상이다. 그림 11의 결과 영상으로부터 제안하는 기법을 통해 생성한 가상 시점 영상이 DIBR 기법[4]에 의해 생성된 가상 시점 영상보다 경계 잡음 및 객체 왜곡의 측면에서 우수함을 알 수 있다. 그림 11(b)의 빨강색과 파랑색으로 표시한 부분을 보면, 기존 기법의 경우, 가상 시점 영상에서 경계 잡음이 남아있는 것을 확인할 수 있다. 또한 노란색으로 표시한 부분에서는 왜곡이 발생한 것을 확인할 수 있다. 반면에,

그림 10. 생성된 다시점 영상 ('BeerGarden' 시퀀스의 첫 번째 프레임으로 생성한 0~9 시점)

Fig. 10. Generated multi-view images (0~9th view in first frame of 'BeerGarden' sequence)

그림 11. 생성된 가상 시점 영상('Cafe' 시퀀스, 5번 시점) (a) 제안된 알고리즘 (b) DIBR 알고리즘[4]
Fig. 11. Generated virtual view image('Cafe' sequence, 5th view) (a) proposed algorithm (b) DIBR algorithm

그림 11(a)의 제안한 기법에 의한 결과에서는 경계 잡음과 객체 왜곡이 감소하였음을 확인할 수 있다.

한편, 그림 12는 'BeerGarden' 시퀀스에 대한 실험 결과를 부분적으로 확대한 것으로, 왼쪽부터 원본 좌영상, 기존의 IDW 기법[5]을 통해 생성된 가상 시점 영상 그리고 제안하는 기법으로 생성한 가상 시점 영상을 나타내며, 가상

시점 영상은 총 10개의 시점 중 네 번째 시점 영상이다. 그림 12(b)의 빨간색 원으로 표시한 부분을 보면 그림 12(a)에 서와는 달리, 얼굴 영역의 크기가 커지는 객체의 왜곡이 나타난다. 반면에 그림 12(c)에서는 이러한 왜곡이 많이 감소된 것을 확인할 수 있다.

그림 13은 'Samgyetang'시퀀스의 가상 시점 영상 생성 결

그림 12. 'BeerGarden' 시퀀스의 부분 확대된 영상 (a) 원본 좌 영상 (b) 기존 IDW 알고리즘으로 생성한 4번 시점 영상[5] (c) 제안된 알고리즘으로 생성한 4번 시점 영상
Fig. 12. Partially zoomed image of 'BeerGarden' sequence (a) original left image (b) 4th view generated by the existing IDW algorithm[4] (c) 4th view generated by the proposed algorithm

그림 13. 'Samgyetang' 시퀀스의 부분 확대된 영상 (a) 기존 IDW 알고리즘으로 생성한 영상 (b) 제안된 알고리즘으로 생성한 영상
Fig. 13. Partially zoomed image of 'Samgyetang' sequence (a) generated by the existing IDW algorithm (b) generated by the proposed algorithm

과를 부분적으로 확대하여 나타낸 그림이다. 기존의 IDW 기법과 제안된 기법으로 생성한 가상 시점 영상으로, 총 10 개의 시점 중 5번 시점의 영상이다. 빨강색 원으로 표시한 부분에서, 제안된 기법으로 생성한 가상 시점 영상에서 객체의 왜곡이 줄어들었음을 확인할 수 있다.

다음으로 기법의 성능을 보다 객관적으로 평가하기 위해 ITU-R 권고 이중 자극 연속 품질 척도법인 DSCQS(double stimulus continuous quality scale)방법을 이용하여 주관적 화질 평가를 수행하였다^[21]. 평가 대상은 일반 평가자 20명이며, 깊이 지도가 필요한 DIBR 기법^[4]의 특성을 고려하여 'Cafe', 'BeerGarden' 시퀀스를 사용하여 DIBR 기법^[4]과의 주관적 화질을 평가를 수행하였고, 'Samgyetang' 시퀀스를 추가하여 기존의 IDW 기법^[5]과의 평가를 수행하였다. 평가는 입체감, 눈의 피로도 그리고 영상 품질의 세 가지 항목으로 진행하였다. 평가 점수는 각 항목별 5점 만점으로 입체감이 좋을수록, 눈의 피로도가 적을수록, 영상의 품질이

표 1. DIBR 및 제안된 기법에 의해 생성된 다시점 영상의 DSCQS 주관적 화질 평가 결과
Table 1. Result of DSCQS subjective quality test for multi-view image generated by DIBR and proposed algorithm

평가기준 \ 기법	DIBR 기법[4]	제안된 기법
입 체 감	3.61	4.13
눈의 피로도	3.16	3.67
영상 품질	3.34	3.87
평 균	3.37	3.89

표 2. 기존 IDW 및 제안된 기법에 의해 생성된 다시점 영상의 DSCQS 주관적 화질 평가 결과
Table 2. Result of DSCQS subjective quality test for multi-view image generated by existing IDW and proposed algorithm

평가기준 \ 기법	기존 IDW 기법 ^[5]	제안된 기법
입 체 감	3.93	4.22
눈의 피로도	3.38	3.51
영상 품질	3.57	3.83
평 균	3.62	3.85

좋을수록 높은 점수를 주도록 하였다. 각 기법을 통해 생성된 다시점 입체 영상에 대하여 주관적 화질을 평가한 후, 최종적으로 각 기법별, 항목별로 평가된 점수의 평균값을 취하였다. 표 1은 DIBR 기법^[4]과의 주관적 화질 평가의 결과를 나타낸 것이고, 표 2는 기존 IDW 기법^[5]과의 주관적 화질 평가의 결과를 나타낸 것이다.

표 1과 표 2를 보면, 입체감, 눈의 피로도 그리고 영상 품질의 항목에 있어서 DIBR 및 기존의 IDW 기법에 비해 높은 점수를 받은 것을 확인할 수 있다. DIBR 기법과 기존 IDW 기법의 경우, 그림 11과 그림 12, 그림 13에서 보인 객체의 왜곡이나 경계 잡음들로 인하여 위와 같은 평가를 받은 것으로 분석된다. 이는 제안된 기법의 가상 시점에서의 객체의 왜곡 감소가 주관적 화질 향상에 좋은 성능을 보인다는 것을 확인시켜준다.

표 3은 제안된 기법의 항목별 수행 시간을 나타낸다. 'Cafe', 'BeerGarden', 'Samgyetang' 시퀀스의 각 100 프

레이를 사용하여 주어진 실험 환경(Visual Studio 2010, Windows 7 64 bit, 8GB Ram, Core i7-2600K 3.4 GHz)에서 실험을 진행하였으며, 평균 시간을 나타내었다. 표 3을 보면, 스테레오 영상으로부터 특징 성분들의 추출 과정과 최적화 과정에 비하여 가상 시점 영상을 생성하는데 상대적으로 적은 시간이 걸리는 것을 확인할 수 있다.

표 3. 제안된 기법의 수행 시간
Table 3. Processing time of proposed algorithm

	수행 시간
변이 특징 지도 추출	16832 ms
명암 변화 특징 지도 추출	883 ms
직선 성분 검출	1411 ms
시간적 불연속성 정제	751 ms
최적화 과정	10345 ms
가상 시점 영상 생성(10개 시점)	1871 ms

IV. 결 론

본 논문에서는 스테레오 영상으로부터 다양한 특징을 추출하고, 차폐영역 추출정보를 입력받아 그리드 메쉬 기반의 워핑 기법에 의해 다시점 영상을 생성하는 기법을 제안하였다. 제안된 기법에서는 모든 화소가 아닌 주요 특징점들에 대한 변이만을 활용하여 다시점 영상을 생성하며, 추출된 특징에 의해 정의된 제한 조건을 사용하여 가상 시점 영상에서 발생하는 왜곡을 최소화할 수 있다. 또한 차폐영역 추출 정보를 활용하여 객체의 경계 영역에서의 변이 정확도를 높여 부정확한 변이로 인한 영상의 왜곡을 줄일 수 있었다. 실험 결과와 주관적 화질 평가를 통해 제안된 기법이 DIBR 기법 및 기존의 IDW 기법에 비해 향상된 주관적인 화질을 보이는 것을 알 수 있었다. 특히 제안한 기법이 객체 경계 영역에서의 왜곡을 줄이고 경계 잡음도 감소시키는 결과를 보이는데, 이는 다시점 영상 기반의 입체 영상에서 눈의 피로도 감소 및 입체감 향상에도 영향을 줄 수 있다.

또한 100 프레임 영상에서 10개의 가상 시점 영상을 생성하는데 평균적으로 1.8초가 소요되는데, 병렬프로그래밍

이나 GPU를 활용하여 처리속도를 개선하고, 최적화 과정의 결과를 스테레오 영상과 함께 부가데이터로 전송하는 방식으로 활용한다면 실시간 처리에 이용할 수 있을 것으로 생각된다.

참 고 문 헌(References)

- [1] B. Bartczak and R. Koch, "Dense depth maps from low resolution time-of-flight depth and high resolution color views," Proc. of 5th International Symposium on Visual Computing, pp.1-12, Nov. 2009.
- [2] ISO/IEC JTC1/SC29/WG11, Draft call for proposals on 3D video coding technology, N11830, Daegu, Korea, Jan. 2011.
- [3] ISO/IEC JTC1/SC29/WG11, Applications and requirements on 3D video coding, N11829, Daegu, Korea, Jan. 2011.
- [4] VSRS: http://wg11.sc29.org/svn/repos/MPEG-4/test/trunk/3D/view_synthesis/VSRS
- [5] W. R. Lee, M. S. Ko, K. M. Um, W. S. Cheong, N. Hur and J. Yoo, "Multi-view image generation from stereoscopic image features and the occlusion region extraction", The Journal of Korean Society of Broadcast Engineers, Vol. 17, No. 5, pp. 838-850, Sep. 2012
- [6] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and Gross, "Non-linear disparity mapping for stereoscopic 3D," ACM Transactions on Graph(SIGGRAPH 2010), vol. 29, July 2010.
- [7] ISO/IEC JTC1/SC29/WG11, Boundary noise removal and common hole filling method for VSRS 3.5, M19356, Daegu, Korea, Jan. 2011.
- [8] ISO/IEC JTC1/SC29/WG11, Image domain warping as alternative to DIBR for advanced 3DV applications, M19995, Geneva, Switzerland, March 2011.
- [9] B. D. Lucas, and T. Kanade, "An iterative image registration technique with an application to stereo vision", Proc. of the 1981 DARPA Imaging Understanding Workshop, pp. 121-130, 1981.
- [10] David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision(IJCV), vol. 60, pp.91-110, Nov. 2004
- [11] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk, "Salient region detection and segmentation," International Conference on Computer Vision Systems, vol. 5008, pp.66 - 75, 2008.
- [12] Nguyen Cao, Truong Hai, D. Y. Kim and H. R. Park, "Obtaining Object by Using Optimal Threshold for Saliency Map Thresholding", The Journal of Korea Contents Association, Vol. 11, No. 6, pp.18-25, Jun. 2011.
- [13] S. Montabone and A. Soto, "Human detection using a mobile platform and novel features derived from a visual saliency mechanism," Image and Vision Computing, vol. 28, no. 3, pp. 391 - 402, 2010.
- [14] O. Barinova, V. Lempitsky, E. Tretyak, and P. Kohli, "Geometric image parsing in man-made environments," in ECCV, 2010.
- [15] R.G. von Gioi, J. Jakubowicz, J. M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. PAMI 32, no.4, pp.

- 722-732, Apr. 2010.
- [16] A. Frick, B. Bartczack and B. Koch, "3D-TV LDV content generation with a hybrid tof-multicamera rig", 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video, pp. 1-4, June 2010
- [17] Sang-Beom Lee and Yo-Sung Ho, "Multi-view depth map estimation enhancing temporal consistency", International Technical Conference on Circuits Systems, Computer and Communications, pp. 29-32, July 2008
- [18] Che-han Chang, Chia-Kai Liang, and Yung-Yu Chuang, "Content-aware display adaptation and interactive editing for stereoscopic images." IEEE Transactions on Multimedia, vol. 13, no. 4, pp.589-601, Aug. 2011.
- [19] J. H. Park and H. W. Park, "A mesh-based disparity representation method for view interpolation and stereo image compression," IEEE Transaction on Image Processing, vol.15, no. 7, pp.1751-1762. July 2006.
- [20] Ilkwon Park and Hyeran Byun, "Efficient data representation of stereo images using edge-based mesh optimization," Journal of Broadcast Engineering, vol. 14, no. 3, pp.322-331, May 2009.
- [21] Methodology for subjective assessment of the quality of television picture, ITU-R Recommendation BT.500-11

저 자 소 개

임 종 명

- 2012년 : 광운대학교 전자공학과 학사
- 2012년 ~ 현재 : 광운대학교 전자공학과 석사과정
- 주관심분야 : super-resolution, 3D 입체 영상 처리, 영상 신호 처리

엄 기 문

- 1991년 : 서강대학교 전자공학과(공학사)
- 1993년 : 서강대학교 전자공학과(공학석사)
- 1998년 : 서강대학교 전자공학과(공학박사)
- 2000년 ~ 현재 : 한국전자통신연구원 책임연구원
- 2001년 ~ 2002년 캐나다 CRC(Communications Research Center) 방문연구원
- 주관심분야 : 컴퓨터 비전, 다시점 3DTV, 휴먼팩터

신 흥 창

- 2005년 : 세종대학교 컴퓨터공학과 학사
- 2008년 : 한양대학교 전자통신컴퓨터공학과 석사
- 2009년 ~ : 한국전자통신연구원 연구원
- 주관심분야 : 3차원 영상처리, 3DTV, 실감방송, 영상 합성, GPGPU 등

이 광 순

- 1993년 : 경북대학교 전자공학과 학사
- 1995년 : 경북대학교 전자공학과 석사
- 2004년 : 경북대학교 전자공학과 박사
- 2001년 ~ 현재 : 한국전자통신연구원 입체방송연구실 실장
- 주관심분야 : 3DTV 방송시스템, 무안경 3D, 영상처리 등

저 자 소 개

허 남 호

- 1992년 2월 : 포항공과대학교 전자전기공학과 공학사
- 1994년 2월 : 포항공과대학교 대학원 전자전기공학과 공학석사
- 2000년 2월 : 포항공과대학교 대학원 전자전기공학과 공학박사
- 2000년 4월~현재 : 한국전자통신연구원 방송통신미디어연구부문 방송시스템연구부 부장
- 주관심분야 : 디지털방송시스템 및 차세대방송 기술

유 지 상

- 1985년 : 서울대학교 전자공학과 학사
- 1987년 : 서울대학교 전자공학과 석사
- 1993년 : Purdue University, EE, Ph.D.
- 1997년 ~ 현재 : 광운대학교 전자공학과 교수
- 주관심분야 : 3D 입체 영상처리 / 압축, 웨이블릿 기반 영상처리, 비선형 디지털 신호처리