

## 가상 시점 생성을 위한 기하학적 제약조건을 이용한 예제 기반 영상 인페인팅

최성환 함범섭 손광훈

연세대학교 전기전자공학과

khsohn@yonsei.ac.kr

## Exemplar-based Image Inpainting with Geometric Constraints for View Synthesis

Choi, Sunghwan Ham, Bumsub Sohn, Kwanghoon

School of Electrical and Electronic Engineering, Yonsei University

## 요약

본 논문에서는 가상 시점 생성 시 발생하는 홀 영역의 메움 기법에 관하여 기하학적 제약조건을 이용한 예제 기반 영상 인페인팅 기술을 제안한다. 생성된 가상 시점이 실제 카메라들의 중간에 위치하는 경우, 이 때 발생하는 홀 영역은 시점 간 보상으로 메울 수 있다. 하지만, 가상 시점이 실제 카메라들의 Field-of-View를 벗어난 경우에는 홀 영역에 대한 정보가 없기 때문에 이를 메우는데 어려운 문제가 따른다. 이러한 문제를 해결하기 위해 제안한 방법은 예제 기반의 영상 인페인팅 기술을 이용하여 홀 영역을 메운다. 제안한 방법은 홀 영역의 기하학적 특성을 제약조건으로 우선순위(Priority)를 계산한다. 최적의 패치를 찾기 위해 컬러 정보뿐만 아니라 깊이 정보(Depth)를 활용한다. 또한, 전경 물체의 텍스처로 인하여 잘못된 패치가 선택되는 문제를 해결하기 위해 적응적으로 패치 크기를 변경한다. 실험 결과에서 제안한 방법은 일반적인 홀 메움 방식과 비교하여 객관적 평가와 주관적 평가 모두에서 시각적으로 우수한 결과를 보여주었다.

## 1. 서론

학계와 산업 전반에 걸친 3차원 영상처리 기술의 활발한 연구에 힘입어 우리는 보다 현실적이며 실감나는 영화, 방송, 그리고 다양한 엔터테인먼트를 주변에서 쉽게 체험하고 있다. 안경을 착용하는 양안 시차를 이용한 스테레오스코픽 방식과 함께 현재는 안경 없이 입체감을 느낄 수 있는 무안경식 Autostereoscopic 방식이 연구되고 있다. 이러한 무안경식 모니터의 성공적인 시장 진입을 위해서는 콘텐츠 제작이 시급하다. 하지만, 대부분의 3D 콘텐츠는 스테레오 방식으로 제작되고 있으며 이러한 문제는 기존의 스테레오 콘텐츠를 무안경식 모니터용으로 변환함으로써 해결할 수 있다.

일반적으로 깊이 영상 기반 영상 합성 기술(Depth Image-Based Rendering)이 스테레오 콘텐츠를 무안경식 모니터용 콘텐츠로 변환하는데 주로 쓰인다[1]. DIBR은 크게 3차원 공간으로의 역 투영(Back-projection), 영상 워핑(Image Warping), 후처리의 순서로 구성된다. 깊이 정보를 바탕으로 영상을 3차원 좌표계로 역 투영 시킨 후, 가상 시점의 위치에 따라 영상을 워핑해준다. 이 과정에서 카메라와 가까운 전경 물체에 가려진 부분이 가상 시점의 위치에 따라 노출되어 보여질 수 있으며 이러한 영역을 홀(Hole)이라고 정의한다. 홀 영역은 보이지 않던 영역으로 이에 해당하는 정보가 없는 상태이므로 이를 메우는 성능은 가상 시점 합성 영상의 화질에 큰 영향을 크게 미친다.

홀 영역을 메우는 방법으로는 일반적으로 보간법(Interpolation)이 많이 쓰인다[2]. 쉽게 구현이 가능하다는 장점으로 널리 쓰이는 반면 눈에 거슬리는 블러링(Blurring) 문제가 발생한다. 반면에 인페인팅(Inpainting) 기술은 블러링 문제를 줄이며 보다 시멘틱한 결과를 출력

한다는 점에서 보간법에 비해 발전된 기술로 영상 복원에서 많이 쓰인다. 인페인팅 기술은 크게 구조 기반(Structure-based) 방식[3]과 예제 기반(Exemplar-based) 방식[4]으로 나뉜다. 구조 기반 방식은 PDE와 Diffusion을 이용하여 홀 주변의 영상 구조를 홀 영역으로 전파시킨다. 따라서 홀 영역의 구조가 복원됨으로써 보간법에 비해 의미있는 결과를 출력하게 된다. 하지만 Diffusion에 기반을 두기 때문에 블러링 문제가 어느 정도 발생하는 문제를 수반한다. 반면에 예제 기반 방식[4]은 패치를 중심으로 홀 주변의 구조와 텍스처를 모두 전파하여 홀을 복원한다. 예제 기반 방식은 크게 홀 메움 순서의 우선순위(Priority) 계산과 최적의 패치를 찾는 것으로 구성된다. 우선순위 계산은 패치 내부의 등광도선(Isophote)의 강도를 나타내는 data term과 패치의 신뢰 정도를 나타내는 confidence term으로 구성된다. 따라서 블러링 문제가 발생하지 않으며 의미 있는 결과를 얻을 수 있다. 하지만, 예제 기반 인페인팅 방식을 가상 시점 합성 시 발생하는 홀 영역에 그대로 적용하면 전경 물체가 홀 영역으로 침투하는 문제가 발생한다. 본 논문에서는 깊이 정보와 홀 영역의 기하학적 특성을 제약조건으로 예제 기반 영상 인페인팅 방식을 제안한다.

본 논문의 2장에서는 가상 시점 생성에서의 홀 영역에 대한 특성과 제안하는 기하학적 제한조건을 다룬다. 3장에서는 제안하는 방법에 대해, 그리고 4장에서는 실험 결과를 다룬다. 마지막으로 5장에서 결론을 맺는다.

## 2. 가상 시점 생성 시 발생하는 홀 영역의 특성

가상 시점 생성 시 발생하는 홀 영역의 형태는 수직성을 갖으며 전경



(a) (b) (c)

그림 1. 전경 텍스처가 홀 영역으로 침투하는 문제: a) 홀 영역, b) 예제 기반 영상 인페인팅[4] 결과, c) Ground Truth

물체의 주변부에 주로 형성되는 특성을 가진다. 예를 들어 가상 시점이 실제 카메라의 왼쪽 방향일 경우, 발생하는 홀 영역은 전경 물체들의 오른쪽 면에 위치하며 수직적인 형태를 가진다. 그림 1과 같이 예제 기반 영상 인페인팅 [4]을 그대로 홀 영역에 적용하면 전경 텍스처가 홀 영역으로 침투하는 문제가 발생한다. 하지만, 이러한 수직적인 기하특성을 기반으로 전경에 해당하는 면에 Zero-priority를 할당하면 전경 텍스처가 홀 영역으로 침투하는 문제를 해결할 수 있다.

홀 영역 주변부의 깊이 정보는 전경에 해당하는 영역에서 큰 변이 값을 가지고 배경에 해당하는 영역에서는 작은 변이 값을 가지는 분포를 나타내는 특성이 있다. 깊이가 모호한 부분에서 홀 메움 시 outlier가 발생할 확률이 높기 때문에 이를 이용하여 깊이가 평활한 영역을 우선적으로 처리할 수 있다.

이러한 기하학적, 정량적 특성을 예제 기반 인페인팅의 우선순위 계산 시 제약조건으로 설계하는 방법을 다음 장에서 기술한다.

### 3. 제안하는 방법

#### 3.1. 기하학적 제약조건을 이용한 우선순위 계산

예제 기반 영상 인페인팅은 높은 우선순위에 우선적으로 홀을 메운다는 전략으로 인페인팅을 진행한다 [4]. 우선순위 ( $P(p)$ )에 영향을 미치는 요소는 1) data term ( $D(p)$ )과 2) confidence term ( $C(p)$ ), 그리고 geometric constraint term ( $Z(p)$ )이며 다음과 같이 정의한다.

$$P(p) = C(p)D(p)Z(p) \quad (1)$$

$D(p)$ 는 패치 내부의 보이는 영역 ( $\Psi_p \cap \Phi$ )에서의 등광도선의 강도를 나타내며 다음과 같이 정의한다.

$$D(p) = \frac{|\nabla I_p^\perp \cdot n_p|}{\alpha} \quad (2)$$

식 (2)에서  $\perp$ 는 직교 연산자이며,  $n_p$ 는 점  $p$ 에서의 수직 단위 벡터이다.  $\alpha$ 는 정규화 요소로서 흑백 영상에서 일반적으로 255 값을 가

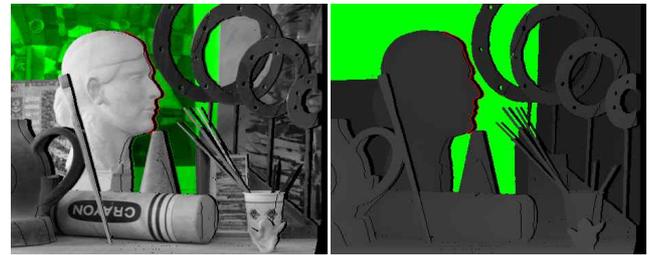


그림 2. 최적의 패치 검색 시 깊이 정보  $Cost_d$ 의 영향으로 현재 홀 메움이 진행 중인 컨투어 주변(빨간색)과 같은 평면에 놓인 녹색 부분에서 낮은 비용함수 값을 가진다.

진다.

$C(p)$ 는 패치의 신뢰도로서 얼마나 많이 유효한(보이는) 픽셀을 가지고 있는지에 대한 비율을 나타내며 다음과 같이 정의한다.

$$C(p) = \left\{ \sum_{q \in \Psi_p \cap \Phi} C(q) \right\} / |\Psi_p| \quad (3)$$

식 (3)에서  $\Psi_p$ 와  $\Phi$ 는 각각  $p$ 점을 중심으로 한 패치, 영상  $I$ 에서 홀이 아닌 보이는 영역을 나타낸다.  $|\Psi_p|$ 는 패치의 크기를 나타낸다.

$Z(p)$ 는 깊이 정보의 분포를 바탕으로 깊이의 모호함을 고려하며 다음과 같이 정의한다.

$$Z(p) = L(p) \left\{ 1 - \frac{\max_{q,r \in \Psi_p \cap \Phi} d(q,r)}{\beta} \right\} \quad (4)$$

식 (4)에서  $d(q,r)$ 는 점  $q$ 와  $r$ 의 깊이 정보의 차이를 나타내며  $\beta$ 는 정규화 값으로서 최대 깊이 값을 가진다.  $L(p)$ 는 이진 함수로써 점  $p$ 가 전경 쪽에 위치할 경우 0을, 배경 쪽에 위치할 경우 1을 가진다.

따라서 우선순위는 1) 패치 내부의 경계 성분 강도가 높을수록, 2) 패치 내부에 보이는 픽셀들이 많을수록, 3) 보이는 픽셀들이 같은 평면에 존재할수록 높은 우선순위를 가진다.

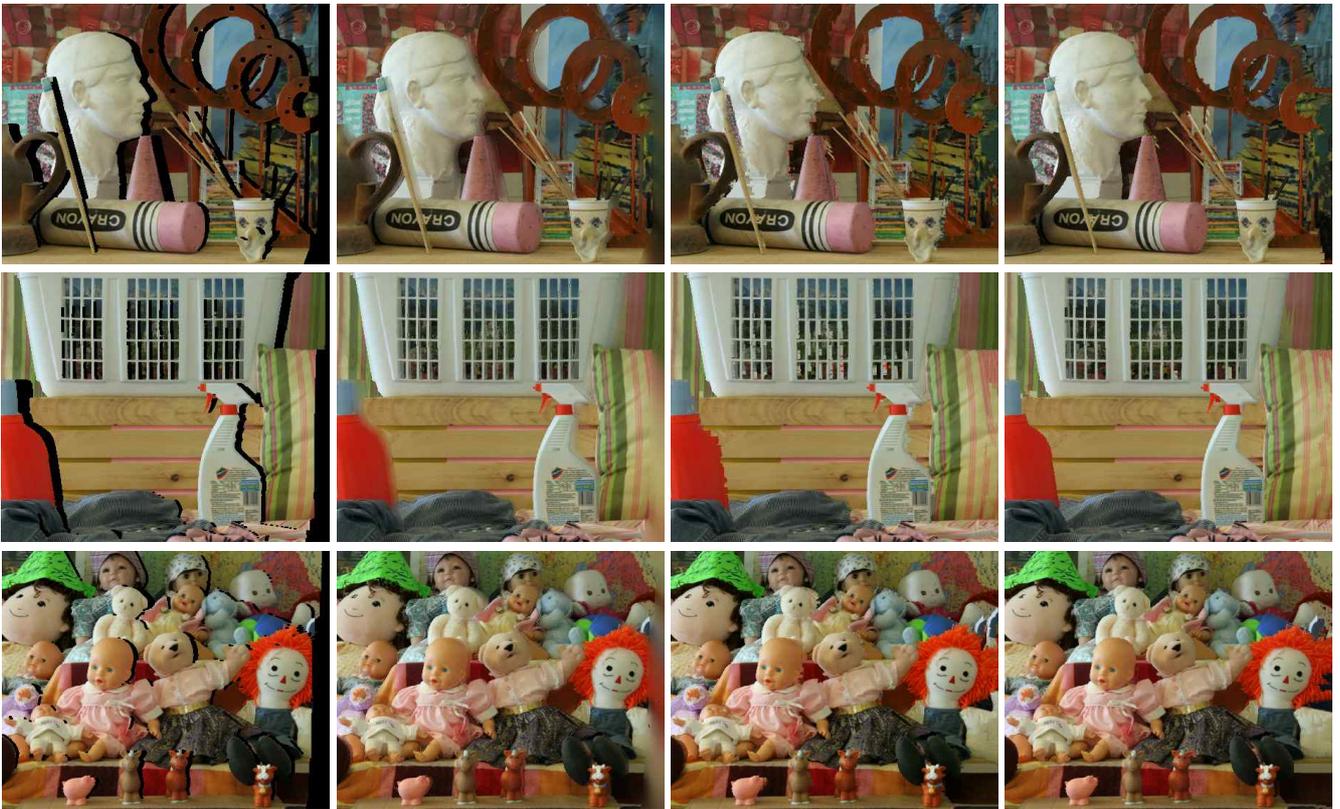
#### 3.2. 최적의 패치 검색

홀 영역을 메울 최적의 패치 후보들은 그 홀 영역과 같은 평면에, 그리고 그 주변부에 있을 확률이 높다. 그 이유는 가상 시점으로 이동하면서 새롭게 노출되는 영역은 그 주변의 연장선상에 위치할 가능성이 높기 때문이다.

최적의 패치 검색은 Sum of Square Difference (SSD) 방식을 이용하여 영상 전체에서 혹은 특정 범위로 한정하여 패치 내부의 보이는 픽셀들( $\Psi_p \cap \Phi$ ) 간의 차이 값으로 구하며, 이 값이 가장 작은 패치를 최적의 패치( $\Psi_p$ )로 선택한다.

$$\Psi_p = \operatorname{argmin}_{\Psi_q \in \Phi} \{ Cost_c(\Psi_p, \Psi_q) + Cost_d(\Psi_p, \Psi_q) \} \quad (5)$$

식 (5)에서  $Cost_c(\Psi_p, \Psi_q)$ 는 컬러 값의 유클리디언 거리를 나타내



(a) (b) (c) (d)

그림 3. 'Art', 'Laundry', 'Dolls' 실험 결과: a) 홀 영역, b) 보간법[2], c) 예제 기반 영상 인페인팅[4], d) 제안한 방법

는 비용함수이며 실험에는 CIELab 색상 공간이 이용되었다.  $Cost_d(\Psi_p, \Psi_q)$ 는 깊이 정보의 차이를 나타내며 다음과 같이 정의한다.

$$Cost_d(\Psi_p, \Psi_q) = \gamma \sum_{r \in \Psi_p, s \in \Psi_q} \min(1, |I_d(r) - I_d(s)|) \quad (6)$$

식 (6)에서  $\gamma$ 는 깊이 비용 함수의 가중치를 나타내며,  $I_d(p)$ 는 깊이 영상을 나타낸다. 그림 2에서  $Cost_d$ 의 영향으로 같은 평면에 위치하는 녹색 부분에서 최적의 패치가 검색됨을 볼 수 있다.

따라서 최적의 패치  $\Psi_p$ 는 1) 컬러 차이가 작을수록, 2) 패치가 같은 평면에 위치할수록 선택되게 된다.

### 3.3. 적응적 패치 크기

패치의 크기를 고정할 경우, 홀 메움이 진행함에 따라 패치가 전경과 배경 모두에 걸쳐 위치하는 경우가 발생하게 되며 이로 인해 경우에 따라 최적의 패치가 전경 영역에서 선택될 가능성이 발생한다. 이 문제는 패치의 크기를 홀 영역의 크기와 비례하여 적응적으로 변경함으로써 패치 내부에 전경 물체가 존재하는 경우를 제거하여 해결할 수 있다.

패치 내부의 전경 텍스처 존재 유무는  $\Psi_p$ 에서의 마스크 값으로부터 알 수 있다.  $\Psi_p$ 에서의 마스크 값에 1차 미분을 x축 방향으로 수행하여 미분 값이 1 이상을 나타낼 때, 패치의 크기를 단계적으로 줄여나

가는 과정을 재귀적으로 반복함으로써 최적의 패치 크기를 얻을 수 있다.

### 4. 실험 결과

실험은 Middlebury 스테레오 데이터[5]를 이용하여 View 1과 그에 해당하는 Ground Truth 깊이 맵을 바탕으로 가상의 View 3을 생성하여 홀 메움을 실행하였다. 모든 실험에는 패치 크기  $9 \times 9$ ,  $\gamma = 60$ 으로 동일한 파라미터가 쓰였다. 실험 결과의 객관적 성능 평가를 위해 PSNR과 SSIM 방식이 이용되었다. 다른 홀 메움 방식과의 비교를 위해 보간법[2]과 예제 기반 인페인팅 방식[4]이 비교에 이용되었다.

그림 3에서 제안한 방식이 다른 알고리즘에 비해 시각적으로 의미 있는 결과를 보여주는 것을 확인할 수 있다. 그림 3-(b)의 보간법에 의한 홀 메움 결과에서 큰 홀에서 눈에 거슬리는 블러링 문제가 발생하는 것을 볼 수 있다. 또한, 그림 3-(c)에서 예제 기반 영상 인페인팅이 블러링 문제는 동반하지 않지만 전경 텍스처가 홀 영역으로 침투하는 것을 볼 수 있다. 하지만, 영상의 우측 큰 홀 영역에서는 홀 메움이 상당히 의미 있는 결과로 채워진 것을 볼 수 있다. 이는 예제 기반 영상 인페인팅이 큰 홀에서 그 성능이 우수하다는 것을 알 수 있다. 그림 3-(d)에서 제안한 방식은 기하학적 제약조건을 바탕으로 텍스처 침투 문제를 해결하였고 큰 홀 영역에서 우수한 결과를 보여주는 것을 확인할 수 있다. 표 1은 객관적 평가로서 PSNR(dB)과 SSIM(Index)을 Ground Truth 영상으로 측정된 결과이다. 표 1에서 제안한 방법이

표 1. Middlebury Stereo Data Set의 객관적 성능평가 결과

	보간법[2]		인페인팅[4]		제안한 방법	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Art	25.02	0.8610	23.03	0.8343	25.47	0.8724
Laundry	26.87	0.8993	24.43	0.8824	27.69	0.9130
Dolls	25.67	0.9043	24.16	0.8859	25.41	0.9268

PSNR과 SSIM 모두 높은 점수를 가지는 것을 볼 수 있다. 하지만 제안한 방법이 시각적으로 더욱 우수하지만 PSNR 측정에서 보간법과 큰 차이가 나지 않는 것은 PSNR은 영상의 차이를 바탕으로 측정하므로 보간법이 PSNR 측정에 유리하게 작용하기 때문이다.

## 5. 결론

본 논문에서는 가상 시점 생성 시 발생하는 홀 영역에 대하여 기하학적 제한조건을 이용한 예제 기반 영상 인페인팅 기술을 제안하였다. 제안한 방식은 홀 영역의 기하학적, 양적 특성을 바탕으로 제약조건을 설계하였고 이를 홀 매움 순서의 우선순위 계산에 적용함으로써 전경 물체의 텍스처가 홀 영역으로 침투하는 문제를 해결하고 깊이의 모호함이 존재하는 부분을 고려하였다. 또한, 최적의 패치 검색 시 깊이 정보를 활용함으로써 결과의 정확성을 높였다. 실험 결과에서 제안한 방식이 비교 알고리즘과 비교하여 우수한 결과를 보임을 확인하였다. 차후 연구로는 시점 간 그리고 시간축 일관성을 고려한 비디오 인페인팅 기술로 확장할 계획이다.

## 참고 문헌

- [1] C. Fehn, "Depth-image-based rendering (dibr), compression and transmission for a new approach on 3d-tv," *in Proc. SPIE Conf. Stereoscopic Displays and Virtual Reality Systems XI*, Jan 2004.
- [2] S. E. Chen, and L. Williams, "View Interpolation for Image Synthesis," *ACM SIGGRAPH*, pp. 279-288, Aug. 1993.
- [3] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image Inpainting," *ACM SIGGRAPH*, pp. 417-424, Jul. 2000.
- [4] A. Criminisi, P. Perez, and K. Toyama, "Region Filling and Object Removal by Exemplar-Based Image Inpainting," *IEEE Trans. Image Processing*, Vol. 13, No. 9, pp. 1200-1212, Sep. 2004.
- [5] <http://vision.middlebury.edu/stereo>