

# 트라이맵 공유를 이용한 다시점 영상의 매팅 및 합성

현명한, 김성열, 호요성  
 광주과학기술원 정보통신공학과  
 {mhhyun, sykim75, hoyo}@gist.ac.kr

## Multi-view Image Matting and Compositing

### Using Trimap Sharing

Myung-Han Hyun, Sung-Yeol Kim, Yo-Sung Ho  
 Gwangju Institute of Science and Technology(GIST)

#### 요 약

디지털 매팅(digital matting)은 원 영상으로부터 전경 객체를 분리하여 특수 효과를 제작하는 기술이다. 최근에 다시점 영상으로부터 생성된 합성 개구 영상(synthetic aperture image)을 사용하여 알파맷(alpha matte)을 생성하는 디지털 매팅 알고리즘이 개발되었다. 그러나 기존의 방법은 다시점 카메라를 사용함에도 불구하고, 오직 단일 시점의 알파맷만을 만들어 다시점 카메라 시스템의 장점을 살리지 못했다. 본 논문은 기존 시점의 트라이맵(trimap)을 모든 시점과 공유하여, 시점에 종속적인 알파맷들을 만들어 모든 시점의 전경 객체들을 효율적으로 추출하는 방법을 제안한다. 또한, 제안한 방법은 모든 시점의 전경 객체를 같은 카메라 구조에서 촬영한 다시점 배경과 합성하여 다시점 합성영상을 만든다.

최근 방송과 영화 산업에서 특수 효과 제작을 위한 매트 및 비디오 합성 기술의 수요가 커지고 있다. 영상 합성 기술은 프로그램 제작비용을 최소화하여 경제적으로 고품질 방송 및 영화 콘텐츠를 만드는 데에 필수적이다. 배경 영상의 전경 객체를 추출하는 기술을 디지털 매팅(digital matting)이라 하고, 추출한 전경 객체를 임의의 배경에 합성하는 기술을 영상 합성(digital compositing)이라 한다 [1][2]. 기본적으로 영상 합성 기술은 크게 스튜디오 영상 합성(studio matting and compositing) 기술과 실사 영상 합성(natural image matting and compositing) 기술로 나뉜다. 영상 합성 기술을 사용하면 비용이 저렴해졌지만, 그 기술에는 여전히 한계점이 있다. 예를 들어 블루스크린 이용되는 스튜디오 영상 합성 방법은 블루스크린 매팅(bluescreen matting)이다 [3]. 파랑색 또는 녹색으로 제한된 배경은 전경 객체를 추출하는데 용이하지만, 그러나 전경 객체가 배경과 같은 색일 때에는 전경 객체를 추출에 어려움이 있을 뿐 만 아니라, 구현을 위해 가상 스튜디오(virtual studio)와 같은 환경이 필요하다. 실사 영상 합성 방법은 촬영의 제약은 감소하나, 사용 장비와 실행 시간이 오래 걸리고[4], 복잡한 객체의 경우 배경에서는 좋은 성능을 기대하기 어렵다 [5][6]. 최근에 다시점 카메라 시스템을 사용하여 실행 시간이 짧고 자동적으로 수행되는 영상 합성 기술이 개발되었지만, 알파 값 추정에 사용한 확장된 분산 추정

(extended variance alpha estimation)으로는 다시점의 알파맷(alpha matte)을 만들어 낼 수 없었다. 결과적으로, 기존의 방법으로는 다시점 카메라 시스템으로부터 다시점 합성영상을 생성할 수 없다.

본 논문에서 제안하는 방법은 먼저 다시점 영상으로부터 각 시점의 전경 객체를 분리하기 위해, 기존 시점의 트라이맵(trimap)을 다른 모든 시점에서 공유한다. 그런 다음, 미지영역 정보를 이용하여 각 시점에 종속적인 알파맷들을 만든다. 마지막으로 동일한 카메라 구조에서 찍은 다시점 배경과 그에 상응하는 분리된 다시점 전경 객체들을 합성하여 다시점 합성영상을 생성한다.

## 2. 제안하는 다시점 영상 매팅 및 합성

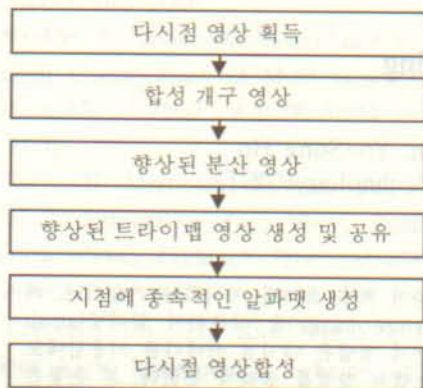
삼차원 영상 편집 시스템에서 다시점 합성영상을 생성하기 위해, 모든 시점의 전경 객체 분리 방법과 합성할 배경에 대한 고려가 필요하다. 본 논문에서는 다시점 영상 매팅 및 합성에 관한 개념과 방법론을 제안한다.



(그림 318) 다시점 영상 매팅 및 합성 방법의 개요

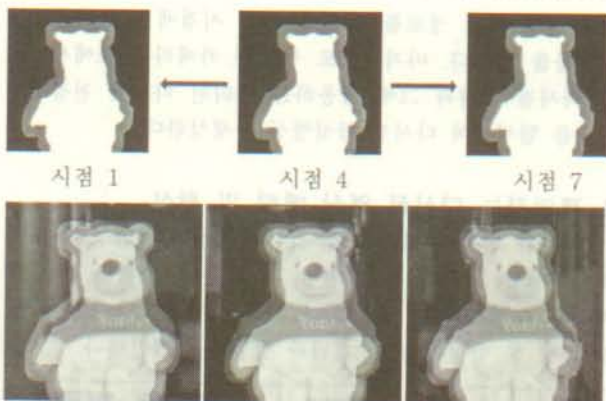


그림 1은 제안하는 다시점 영상 매칭 및 합성 방법의 개요를 보여준다. 먼저 다시점 카메라 시스템으로부터 다시점 영상을 촬영하고, 다시점 영상의 전경 객체들을 분리한다. 분리된 전경 객체들을 같은 카메라 구조에서 찍은 다시점 배경 영상과 합성한다. 다시점 카메라를 이용한 디지털 매칭의 장점은 사용자의 개입 없이 트라이앵플을 만들 수 있다는 점이다. 기존의 방법은 기존 시점의 전경 객체만을 분리할 수 있었지만, 제안하는 방법은 모든 시점의 전경 객체들을 분리해 낼 수 있다.



(그림 319) 다시점 영상 매칭 및 합성 방법의 흐름도

그림 2의 흐름도는 제안하는 다시점 영상 매칭 및 합성 방법을 상세히 설명한다. 다시점 영상으로부터 합성 개구 영상, 분산영상, 트라이앵플을 만들고, 생성된 트라이앵플을 모든 시점과 공유하여 시점에 종속적인 알파맵을 만든다. 이 알파맵으로 분리한 모든 전경 객체들을 다시점 배경과 합성해서 다시점 합성영상을 만든다. 결과적으로, 제안하는 방법은 모든 시점의 트라이앵플을 수작업으로 만드는 수고를 줄인다.

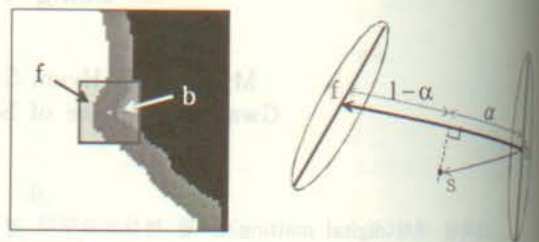


(그림 3) 트라이앵플 공유 및 포함된 경계

그림 3에서 보듯이, 시점 4의 트라이앵플은 모든 시점 객체의 경계부분을 포함하게 된다. 이 경계 부분의 정보가 시점에 종속적인 다시점 알파맵을 만드는 데에 사용된다.

기존의 방법은 모든 시점의 분산 영상을 만들 수 없기 때문에 기존 시점을 제외한 시점의 알파맵들을 만들기가

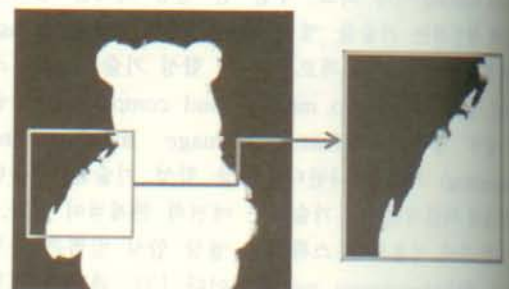
어렵다. 제안하는 방법은 기존 시점의 트라이앵플을 모든 시점과 공유하여 모든 시점의 알파맵들을 추정할 수 있다. 그림 4(a)는 미지 영역의 한 점을 기준으로 윈도우를 형성함으로써 추출된 전경, 배경의 화소 집합을 보여준다. 미지 영역의 넓이에 따라 윈도우의 크기를 적절하게 바꾼다. 미지 영역의 알파 값을 구하기 위해서, 미지 영역의 화소들과 그에 대응하는 전경 및 배경의 대표 점을 구한다. 논문에서는 R, G, B 좌표를 레벨 값 10으로 양자화하고 그 값의 히스토그램을 만든다. 여기서 히스토그램의 빈도가 적은 것은 잡음이라 가정하고, 크기가 10 이상인 것만 가지고 평균을 취하여 대표 R, G, B 좌표를 얻는다.



(a) 화소 집합 (b) 색상 공간  
(그림 4) 색상 집합과 색상 공간

전경, 배경, 미지 영역의 대표점들을 식 (1)에 대입하여 불투명도  $\alpha$  값을 구한다. 즉, 그림 4(b)의 벡터  $\vec{f}-\vec{b}$ 와  $\vec{s}-\vec{b}$ 를 내적하여  $\alpha$  값을 구한다. 여기서  $\vec{f}$ 와  $\vec{b}$ 는 전경과 배경의 대표점을 나타내는 벡터이고,  $\vec{s}$ 는 미지 영역 화소를 나타내는 벡터이다.

$$\alpha = \frac{(\vec{s}-\vec{b}) \cdot (\vec{f}-\vec{b})}{|\vec{f}-\vec{b}|^2} \quad (1)$$



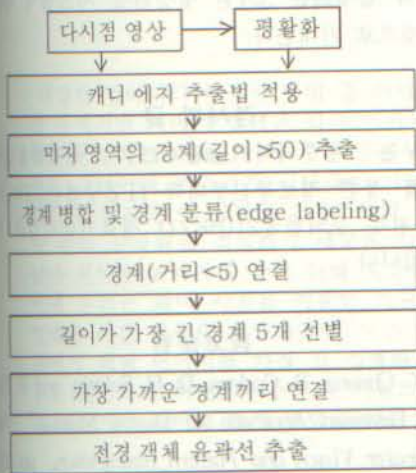
(그림 5) 알파맵 추정의 문제점

디지털 매칭 기술에서 트라이앵플을 생성하는 작업은 가장 어려운 부분에 속한다. 보통 트라이앵플은 사용자의 주관적인 개입으로 만들어지고, 이것은 정확한 알파맵을 구하는 데에 가장 큰 영향을 준다. 비록 제안하는 방법은 다시점 카메라를 사용한 트라이앵플 공유를 통해 손쉽게 모든 시점의 알파맵을 만들 수 있지만, 그림 5에서 보듯이 색상 집단에 기반한 알파 추정 방법을 사용할 때, 객체 전경 부분이 잘 표현되지 못하는 단점이 있다. 또한 객체



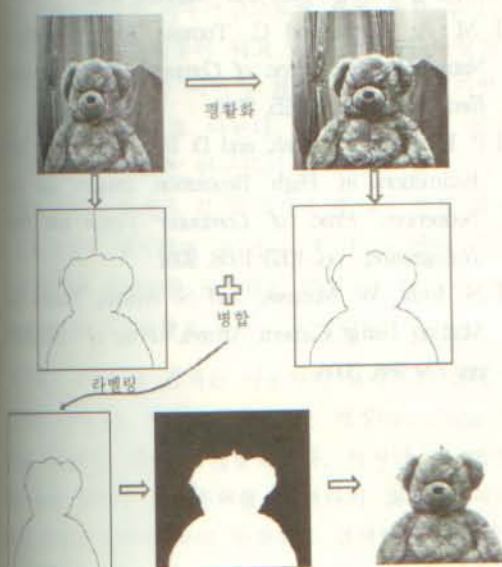
이 기반한 알파 추정 방법은 트라이앵글 미지 영역의 추출을 추정하는 데에 많은 시간이 걸린다.

이 방법의 기반한 알파 추정의 문제점을 해결하기 위하여는 공유한 트라이앵글의 미지 영역 정보를 이용한 그림 6은 전경 객체의 경계를 추출하는 방법의 블록 다이어그램을 나타낸다. 원 영상의 경계를 부각시키기 위해 평활화 및 필터링을 한 후, 트라이앵글 미지 영역 정보를 이용하여 원영상의 경계를 찾아낸다. 찾아 낸 경계 정보를 이용하여 임계치 보다 경계 사이의 길이가 작은 경계는 있는 경계 병합 및 라벨링 과정을 거친 후, 가장 두꺼운 경계를 5개를 찾고 그 경계들을 연결하여 최종적으로 전경 객체의 경계를 찾는다.



(그림 6) 전경 객체 경계 추출 블록 다이어그램

그림 7은 전경 객체 경계 추출과정을 그림으로 보여준 원 영상과 평활화 영상의 경계 병합 및 라벨링 과정과 최종 경계를 추출하고, 그것을 기반으로 알파맵을 만들어 최종 전경 객체를 추출하는 과정을 보여준다.



(그림 7) 전경 객체 경계 추출 과정

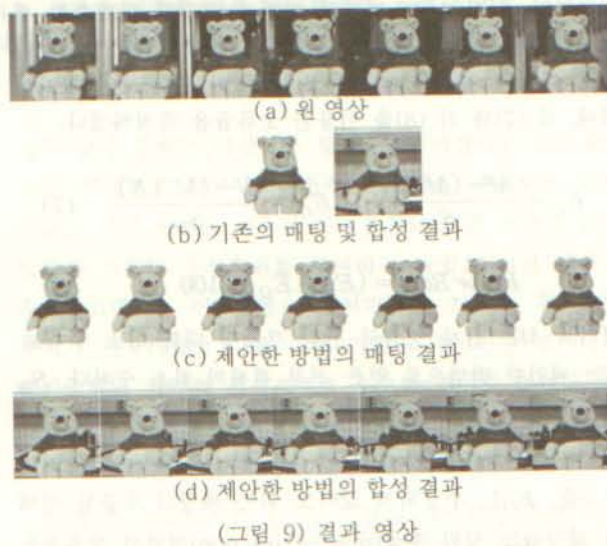
### 3. 실험 결과 및 고찰

본 논문에서는 트라이앵글 공유를 이용한 다시점 영상 매칭 및 합성 방법을 제안하였다. 실험을 위해 그림 8과 같이 일차원 평행 구조의 다시점 카메라 7개를 5cm 간격으로 배치하였다.

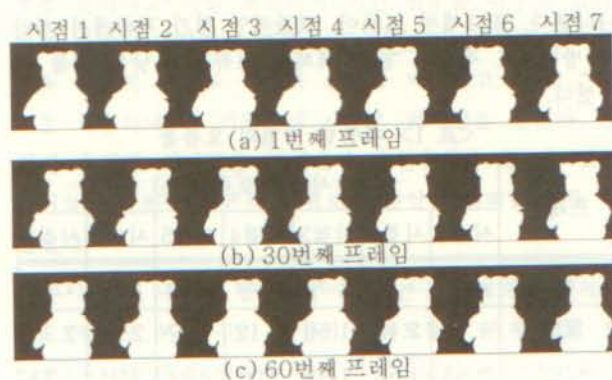


(그림 8) 실험에 사용된 다시점 카메라 시스템

그림 9는 다시점 영상 매칭 및 합성 결과를 보여준다. 제안한 방법으로 다시점 전경 객체를 추출하여 동일한 카메라 구조에서 찍은 다시점 배경과 합성하였다. 그림 9(b)는 다시점 합성영상을 만들 수 없는 기존의 매칭 및 합성 결과를 보여주고, 그림 9(c)와 그림 9(d)는 제안한 방법을 이용하여 다시점 전경 객체들을 다시점 배경과 합성하여 생성한 다시점 합성영상 결과를 보여준다. 그림 10은 트라이앵글 공유를 통하여 만든 시점에 따른 알파맵을 보여준다.



(그림 9) 결과 영상



(그림 10) 시점에 따른 알파맵





(그림 11) 추출된 전경 객체 비교



(그림 12) 알파맵 비교

그림 11은 기존의 방법과 제안한 방법으로 분리한 전경 객체를 비교하는 그림이다. 기존의 방법은 기존 시점의 전경 객체만을 추출하였다. 따라서 기존 시점의 전경 객체만을 가지고 주관적인 비교를 했을 때, 제안한 방법이 경계 부분을 더 잘 표현하였으며 객체 내부의 홀도 제거하였다. 그림 12는 기존의 방법과 제안한 방법으로 생성한 알파맵을 보여주며, 제안한 방법이 기존 방법보다 전경 객체의 경계를 더 잘 표현하였다. 객관적인 평가를 위해, 식 (7)과 식 (8)을 이용한 오류율을 측정하였다.

$$E_U = \frac{M - (M \cap N)}{S_M}, \quad E_O = \frac{N - (M \cap N)}{S_N} \quad (7)$$

$$Error\ Rate = (E_U + E_O) \times 100 \quad (8)$$

여기서  $M$ 은 실제 영상의 전경 객체에 대한 화소 수이고,  $N$ 은 제안한 방법으로 얻은 전경 객체의 화소 수이다.  $S_M$ 과  $S_N$ 은  $M$ 과  $N$ 의 개수를 의미하며,  $E_U$ 는 추출되지 않은 영역에 해당하는 하위 추출(under-extraction)영역의 오류율,  $E_O$ 는 추출되지 않아도 되는 배경이 추출된 영역에 해당하는 상위 추출(over-extraction)영역의 오류율을 나타낸다. 표 1에서 보듯이, 객관적인 평가 기준에서 제안한 방법으로 추출한 전경 객체의 오류율이 낮은 것을 확인했다.

<표 1> 제안한 방법의 오류율

프레임	각 시점의 오류율(%)						
	시점1	시점2	시점3	시점4	시점5	시점6	시점7
1번	2.74	1.73	1.82	1.22	1.35	1.64	1.79
30번	1.13	2.42	1.56	2.12	1.97	2.96	2.36
60번	2.83	1.58	1.95	2.07	2.31	3.01	1.77

#### 4. 결론

본 논문에서는 트라이맵 공유를 이용한 다시점 영상 매칭 및 합성 방법을 제안했다. 제안한 방법은 기존 시점의 트라이맵을 모든 시점과 공유함으로써 모든 시점에 고속적인 알파맵을 만들어냈다. 알파맵 추정을 위해 사용된 색상 집단에 기반한 알파 추정방법을 개선하기 위해 공유된 트라이맵의 미지영역 정보를 이용하여 전경 객체의 경계를 추출해냈다. 추출된 경계를 바탕으로 만들어진 알파맵이 다시점 전경 객체를 효과적으로 분리함을 확인하였고, 객관적인 평가에서도 낮은 오류율을 보였다. 제안한 트라이맵 공유를 이용한 다시점 영상 매칭 및 합성 방법은 다시점 전경 객체를 같은 카메라 구조에서 찍은 다시점 배경과 합성하는 3차원 영상편집 시스템에 유용하게 사용될 것으로 기대한다.

#### 감사의 글

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신부대학 IT연구센터(ITRC)와 교육인적자원부 두뇌한국21(BK21) 정보기술사업의 지원에 의한 것입니다.

#### 참고문헌

- [1] Y. Y. Chuang, B. Curless, D. H. Salesin, and R. Szeliski, "A Bayesian Approach to Digital Matting," *Proc. of Computer Vision and Pattern Recognition*, pp. 264-271, 2001.
- [2] T. Porter and T. Duff, "Compositing Digital Images," *Computer Graphics*, vol. 18, no. 3, pp. 253-259, 1984.
- [3] A. R. Smith and J. F. Blinn, "Blue Screen Matting," *Proc. of SIGGRAPH*, pp. 259-268, 1996.
- [4] P. Hillman and J. Hannah, "Natural Image Matting," *Proc. of Vision, Video and Graphics*, 2005.
- [5] M. A. Ruzon and C. Tomasi, "Alpha Estimation in Natural Images," *Proc. of Computer Vision and Pattern Recognition*, pp. 18-25, 2000.
- [6] P. Hillman, J. Hannah, and D. Renshaw, "Alpha Channel Estimation in High Resolution Images and Image Sequences," *Proc. of Computer Vision and Pattern Recognition*, pp. 1063-1068, 2001.
- [7] N. Joshi, W. Matusik, and S. Avidan, "Natural Video Matting Using Camera Arrays," *Proc. of SIGGRAPH*, pp. 779-785, 2006.