공학박사 학위논문

적응적 트라이맵을 사용한 합성 영상 편집 시스템

2019년 8월

부경대학교 대학원

IT융합응용공학과

윤 요 섭

공학박사 학위논문

적응적 트라이맵을 사용한 합성 영상 편집 시스템

지도교수 김 영 봉

이 논문을 공학박사 학위논문으로 제출함.

2019년 8월

부경대학교 대학원

IT융합응용공학과

윤 요 섭

윤요섭의 공학박사 학위논문을 인준함

2019년 08월 23일(학위수여일)

위원장	공학박사	김 종 남 (인)
위 원	공학박사	신 봉 기 (인)
위 원	이학박사	김 형 석 (인)
위 원	공학박사	조 장 우 (인)
위 원	공학박사	김 영 봉 (인)

목 차

- 표 목차 -

- 그림 목차 -

Blended Video Editor using an Adaptive Tri-map

Joseph Yoon

Dept. of IT Convergence and Application Engineering,
The Graduate School,
Pukyong National University

Abstract

Commercial video blending technologies used in movies, broadcasts, and commercials are complex to utilize by ordinary users because they require expertise. In this paper, we propose an image regeneration system that can get similar image composition results with fewer operations than conventional commercial programs. The proposed system consists of three steps with a video classifier, an image search extractor, and a video blender. Composite image generator is implemented with emphasis on object, face, and background compositing functions. This system is composed of a Poisson Image Matting technique for natural boundary processing of images, a Haar-like feature technique for automatically tracking face positions, sizes and angles, and a Watershed Segmentation technique for easy separation between foreground and background. Especially, we propose a video editor with new blending technique based on an adaptive Tri-map, in order to improve the quality of blended video under various image conditions. The proposed system makes it possible to regenerate high quality blending images through simple manipulation of videos and images synthesis so that it can be easily used by unskilled public

Ⅰ 서론

최근 다양한 형태의 SNS의 보급과 사용자 참여형 TV와 같은 새로운 미디어의 발전으로 인하여, 콘텐츠의 생산에 있어서 대중이 직접 참여할 수 있는 길이 생기게 되었다. 이로 인해 콘텐츠를 제공하는 제작자가 대중의 여론에 부합하는 콘텐츠를 제작·공급하는 것을 넘어서, 사용자 개개인의 생각을 적극적으로 반영할 수 있는 맞춤형 콘텐츠 제작 시스템으로의 전환을 요구받고 있는 실정이다. 이러한 배경에는 최근 ICT 분야의 화두로 떠오른 클라우드와 소셜네트워크 플랫폼 기술들이 보급되기 시작하면서, 소셜커머스, 소셜게임 등의 SNS를 활용한 서비스와 기술 인프라가 구축되어졌기 때문이다. 이는 곧 더 발전적인 형태의 SNS기반 신개념의 융합 기술들의 등장을 초래하였다. 예를 들어, 웹 환경에서의 실시간 댓글이나 스마트 디바이스를 통한 소비자의 선택에 따른 콘텐츠를 제공하는 사용자 참여에 중심을 둔 소셜네트워크 광고, 그리고 멀티 클라이언트 영상 수신 디바이스의 시대적 흐름에 따른 N-스크린 환경에서의 대중 콘텐츠 제작방식인 소셜TV와 같은 새로운 패러다임의 기술들이 각광받고 있다. 즉, 사용자 맞춤형, 사용자 참여형 미디어 콘텐츠들에 대한 요구가 대폭적으로 증가하고 있는 상황이다.

동영상 관련 기술들은 아날로그로 대표되던 필름 영화에서 디지털 영화로 넘어온 지가 벌써 10여년이 지날 만큼 디지털화가 급속히 진행되어 왔다. 동영상의 디지털화는 영상에 대한 각종 처리를 가능하게 해준다. 하나의 동영상에서 배경과 인물을 분리하는 것도 가능하고, 주인공의 움직임을 추적하는 기술도 가능하게 되었다. 최근에 들어서는 동영상으로부터 주인공의 얼굴 모습을 3차원으로 만드는 기술에 대한 연구도 이루어지고 있는 실정이다. 이런 기술들은 각자의 영역에서 개발

이 되어 오다가 최근에는 각 기술들을 사용자의 요구에 맞추어 서로 융합시키는 연구에 대한 관심이 증가하고 있는 실정이다. 그래서 기존의 동영상에서 배경을 원하는 지역이나 장면으로 변경하는 시도도 있고, 주인공의 얼굴을 사용자가 좋아하는 배우의 얼굴로 대체하여 패러디 영상을 만드는 시도가 늘고 있다. 이렇게 사용자가 직접 원하는 콘텐츠를 만드는 작업과 사용자의 참여를 가능하게 해주는 SNS가 연결되어 협업을 시도하는 연구가 매우 중요하게 되었다.

이에 본 연구에서는 미디어 매체를 통해 단방향으로만 제공되어 왔던 영상 콘텐츠를, 사용자의 참여를 바탕으로 한 반응형 콘텐츠로 재생성하는 기법에 대하여 논의하고자 한다. 이는 기존의 텍스트, 오디오, 제작된 UCC, VOD 서비스 등의 콘텐츠 제공 방식에만 국한하지 않고, 한차원 높은 동영상 미디어를 수요자가 직접 콘텐츠를 수정, 제작하고, 사용할 수 있도록 하는 기술 플랫폼에 적용할 수 있는 콘텐츠 제작 기법에 관한 기반 기술을 개발하고자 한다.

이는 스마트기기의 보급 확산과 무선 인터넷 인프라의 발전과 맞물려 고화질 영상 등의 대용량 데이터트래픽 처리가 가능한 현재의 ICT 기술 환경 하에서 차세대 미래 기술을 선점하기 위한 연구로서 반드시 필요할 것으로 생각되어진다. 이를 위한 전초단계로써 각각의 영상에 붙여진 태그를 기반으로 원하는 종류의 영상을 검색하고, 검색된 영상들 중에서 사용자가 원하는 적합한 영상을 선택한 후 선택된 여러 개의 영상을 합성하는 기법에 대한 연구와 영상 합성기를 개발하는 것이 필요하다.

1. 연구배경

현재 사용자 참여 기반의 영상 제공 기술의 경우엔 다양한 형태의 생산 활용 단계의 연구가 진행되고 있다. 영상의 재생성이 아닌 기존의 영상을 소셜 TV라는 형태의 VOD(Video On Demand) 서비스(즉, 사용자가 원할 때 혹은 소셜 네트워킹으로 공유된 친구나 동료들과 원할 때 영상을 제공받는 서비스 형태는 미국에서 상용화된 제품이 존재하고 있다. 현재 다양한 계층의 사용자들의 피드백에 따라 기술 형태를 재정립하고 있다. 다음 그림1의 예에서 보여주듯이 개발 기술 평가는 현재 제공되는 서비스 기술 수준에 기인한다.

그림 1. 소셜 TV 영상 서비스 분석을 통한 기술 수준 측정

영상의 재생성에 필요한 여러 필수 기반 기술들 중에서, 동영상 및 이미지 처리 분야에서 추출(얼굴,객체)과 합성(이미지와 영상) 관련 단위 기술들은 실시간 처리가 가능할 만큼 기술 안정화 단계에 있으나 실제 상용화 형태로는 편집 멀티미디어 저작도구의 플러그인 형태로 제공

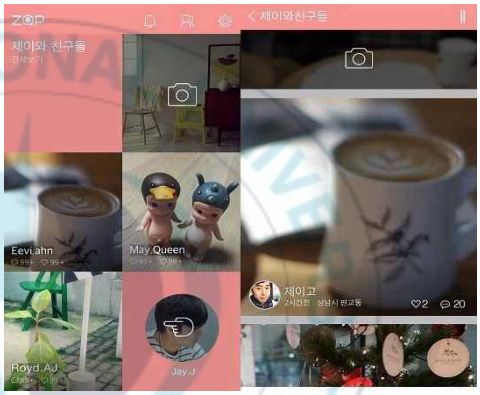
된다, 쉽게 말해 영화나 방송제작자들이 사용하는 소프트웨어 안에 내장된 형태의 기술로 존재한다는 의미다. 아직은 전문가가 아닌 일반 사용자들이 사용하는 도구에서는 해당 기능들이 제공되지 않고 있다. 하지만 SNS기반의 여러 회사들을 필두로 하여 간단한 영상의 수정이 가능한 서비스를 제공하면서 생산·활용단계로 진입하기 위한 시도가 이루어지고 있다.

국내 기술의 경우 네이버와 카카오 같은 인터넷 대형 포털 벤더의 연구소를 중심으로 활발히 기술 연구가 진행되어지고 있으며, 그 결과 실험적인 상용 제품들이 그림2와 같이 출시되었다. 주요 기능을 살펴보면 기본적으로 짧은 분량의 이미지 및 영상 촬영과 공유기능에, 추가적으로 영상의 속도를 다르게 조절한다던가, 필터를 적용하여 색상을 변경하거나 콘텐츠를 꾸미기 위한 배경 이미지나 이모티콘 스티커를 영상에 적용하는 등의 간단한 편집 기능들을 제공하고 있다. 물론 이들은 사용자들의 피드백을 받으며 지속적으로 기능을 수정 보완해 나가고 있다. 또한 공통적인 특징으로는 일반 사용자가 제작한 콘텐츠를 공유하고 재생산하기 위해서 콘텐츠를 데이터베이스화한 플랫폼 개발을 목표로 한다는 것이다. 하지만 이들이 제공하는 영상의 제작, 공유 기능에 비해서, 기존의 영상을 활용하여 편집한 후, 새로운 영상을 재생성하는 기술은 거의 적용이 안 되고 있는 실정이다.

1. 네이버- '폴라(pholar)'

2. 다음카카오 - '쨉(zap)'

3. 네이트 - 싸이메라(cymera)



4. 디콘팩토리 - 폰더(fonder)

그림 2. 사용자 맞춤형 영상 제작과 공유를 위한 국내 상용 서비스

국외 현황을 살펴보면 유튜브와 같은 사용자 제작 동영상 플랫폼으로 콘텐츠 시장이 재편되었다. 요즘 가장 각광받고 있는 소셜네트워크 서비스(SNS) 기술 분야를 중심으로 영상 재성성 기술을 살펴보면, 단순 텍스트에 기반한 트위터와 텍스트, 이미지 공유 중심의 페이스북의 수요가 정체된 반면에, 이미지와 동영상 등의 사용자 관심사에 알맞게 콘텐츠를 큐레이션하는 인스타그램, 틱톡, 텀블러와 같은 기술 서비스들의 수요가 폭발적으로 증가하였다.

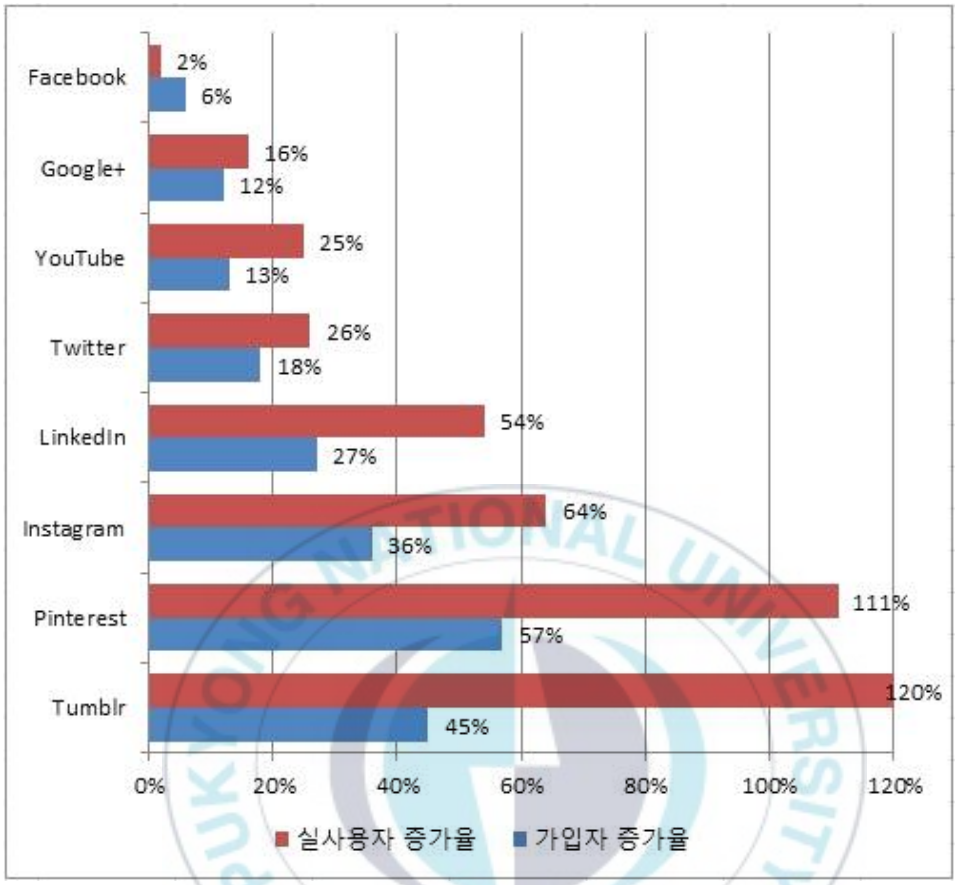


그림 3. 상용 서비스별 가입자와 실사용자 증가율

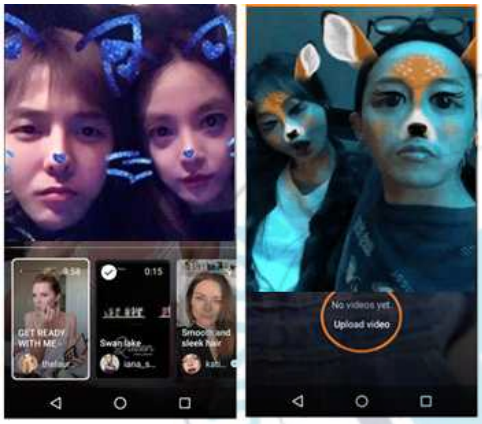
예를들어, 인스터그램의 경우 2019년 현재 시점에서 월간 이용자(MAU) 10억명으로, 특히 비디오 서비스인 스토리는 하루 이용자수가 5억명을 넘어섰다. 참고로 텍스트 전용인 트위터는 이용자수가 3억3천명으로 3배이상의 차이가 난다. 또한 동영상 콘텐츠 전용의 틱톡의 경우 출시 2년만에 5억명을 넘어섰다. 또한 그림3 에서 확인 할 수 있듯이 동영상 중심의 텀블러 또한 실사용자수가 120% 급격하게 증가하였다. (출처:미국 글로벌 웹인덱스). 이는 동영상에 대한 사용자의 수요가 크다고 볼 수 있고, 앞으로도 지속적인 증가가 될 것으로 예상된다.

또 다른 지표로 그림 4와 같이 모바일 앱 어플리케이션의 상위 다운로드 수를 살펴보면, 최상위에 15초 이내의 소셜 동영상을 제작하는 틱톡, 2위에 최대 12시간의 동영상까지 업로드하여 공유할 수 있는 플랫폼인 유튜브, 그 외 최소10분에서 최대 1시간까지의 고화질 동영상을 편집, 업로드 하는 인스타그램의 IGTV가 상위에 랭크되어있다. 이러한 점에서 일반 사용자들의 동영상 공유, 편집 제작 시스템에 대한 기대가 크다고 사료되어진다.

그림 4. 앱 스토어의 어플리케이션 다운로드 순위

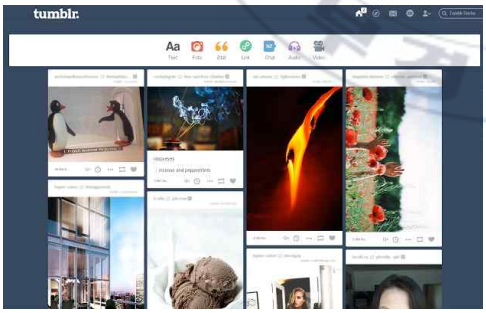
이러한 현상은 기존 텍스트 중심의 콘텐츠들을 화면이 작은 모바일 기기에서 디스플레이 하기 위한 대안으로, 작은 디스플레이의 단점을 보완할 수 있도록, 콘텐츠의 형태가 사진이나 동영상등의 시각물 중심으로 변하고 있기 때문으로 분석된다. 또 하나의 현상은 원하지 않는 정보를 반강제적으로 봐야하는 단점을 보완하기 위해, 철저한 사용자의

관심사에 의해 필터된 맞춤형 콘텐츠를 사용자는 요구하고 있으며, 기술은 이에 맞춰 발빠르게 진화하고 있다. 더 나아가 미래에 요구되어지는 기술은 사용자들이 직접 제작에 참여할 수 있는 동영상 콘텐츠 생성, 수정 시스템으로 전망된다. 이에 본 연구에서도 그림5의 글로벌 상용 서비스 플랫폼과 같이, 동영상 콘텐츠의 재생성 등에 사용자의 요구를 반영할 수 있는 합성 영상 재생성기를 개발하고자 하였다.

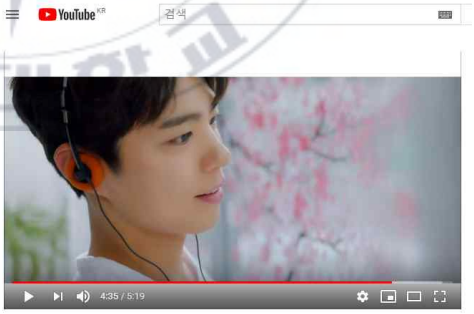


1. 인스타그램TV(IGTV)

2. 틱톡(TikTok)



3. 텀블러(Tumblr)



4. 유튜브(YouTube)

그림 5. 사용자 맞춤형 영상 제작과 공유를 위한 글로벌 상용 서비스

지금까지 서술한 현황을 분석하여 도출한 문제점을 정리하면 아래와 같다.

첫째, 영상 콘텐츠 공급에서 사용자 참여형 서비스가 접목이 되기 위해서는, 기존의 SNS 서비스와 융합도 고려해야 하는데, 이를 위해선 영상의 전체가 아닌 부분적 텍스트 태그 기능을 부여할 수 있는 기능에 대한 연구가 선행되어져야 하는 등의 기술적 난해함이 존재한다.

둘째, 현재 영상처리 편집 기술의 경우 대부분 숙련된 전문가들이 저작 소프트웨어를 통해 영상을 제작할 수밖에 없기 때문에 일반 사용자들이 영상을 재생성하는 것에는 한계가 있다.

셋째, 저작 도구의 경우, 일반 사용자들이 사용하기에는 상당히 고가라는 경제적인 문제점도 지니고 있다. 다시 말해 콘텐츠 제작 기업이나 소프트웨어 개발사의 경우도 영상 재생성의 자유도를 높이기 위해 고가의 비용을 지불해야하는 단점이 있다. 저비용의 오픈소스프로젝트로 개발 환경을 구축할 필요성이 있다.

넷째, 최근 스마트폰, 태블릿, 스마트TV와 같은 다양한 재생 기기들의 확대로 인해 플랫폼에 따른 복잡한 영상처리를 손쉽게 다룰 수 있는 표준화된 사용자 인터페이스를 구현하기 힘들다.

다섯째, 모바일 디바이스의 영상 시청이 현재는 작은 비율을 차지하고 있으나, 사용자의 증가로 트래픽이 심할 경우, 영상 재생성 기법의 계산 처리 속도를 고려해야 한다.

이와 같이 도출된 다양한 문제점을 감안하여, 본 논문에서는 최종적으로 사용자도 손쉽게 맞춤형 콘텐츠를 제작 할 수 있도록하는 영상 재생성 플랫폼을 구현하고자 하였다.

2. 연구내용

본 연구에서는 이미지 검색 추출과 영상 분류 처리 후에 객체, 얼굴, 배경을 구분하여 사용자의 간편 조작만으로도 손쉽게 영상을 재성성할 수 있도록 구현한 영상 편집 시스템에 대해서 소개하고, 특히 구현 모듈 중 영상 합성기에서 적응적 트라이맵을 새로운 객체 합성 기법을 제안한다. 이를 통해 합성 대상 이미지들 간의 다양한 조건에서도 그에 알맞은 합성방법을 선택할 수 있게 하여 최선의 합성 결과를 도출하는 도구를 제공하고자 하였으며, 아울러 최소한의 사용자 조작만으로도 손쉽게 다룰 수 있도록 함으로서, 비숙련자인 일반 사용자들이 합성영상을 제작을 할 때에 편하게 편집 작업을 할 수 있기를 기대한다.

기존의 영화, 방송, 광고 제작에 쓰이는 상용화된 시스템들은 숙련된 전문 기술을 요구하므로, 일반 사용자가 활용하기 힘들다는 문제점이 있다. 이러한 문제를 해결하기 위한 접근 방법으로서 사용자와 시스템간의 인터페이싱 단순화에 중점을 두고 개발하였다. 또한, 인터페이싱 절차의 최적화와 함께 합성시 영상 품질을 유지하고 연산 속도를 보장하기 위하여, 영상안의 대상물을 얼굴, 배경, 객체 등으로 구분하여 각각의 모듈별로 분산시켜 처리하였다.

본 논문에서 중점적으로 논의하고, 구현되어진 시스템의 세부 연구내용은 위의 표1 과 같이 정리된다.

	세부 연구 내용(핵심 요소)
1	원본 영상에서 재생성 할 구간 범위를 설정하는 사용자 인터페이스 개발(예:최소한의 마우스 클릭과 드래그 기능만으로,배경,객체 추출시)
2	원본 영상의 프레임에서 교체될 이미지 영역을 추출하는 기술 개발 (예:적응적 트라이맵 제안- 자동,수동,반자동 선택)
3	변경 영상의 한 프레임에서 대체 이미지의 영역을 선택하는 인터페이스 개발(예:직관적인 그래픽 UI 제공)
4	합성 시, 교체될 이미지 영역에 대체 이미지를 위치시키는 자동화 기술 개발(예:얼굴-특징점 매칭, swapping, 객체-위치 자동화)
5	합성 시, 한 프레임 내의 합성된 두 이미지의 품질을 개선하는 알고리즘 연구(예: 확률, 통계기반의 적응적 트라이맵 기법 제안)
6	재생성 시킬 영상 구간(다수 프레임) 합성 시, 연산 시간을 줄이기 위한 프레임 합성 연산 속도 개선 연구(예:GPU병렬처리 적용)
7	합성 후, 다수 프레임(이미지) 연결 시, 재생성한 전체 영상의 자연스러운 움직임에 관한 연구(예:모션벡터,모핑 함수 산출)

[표 1] 제안 시스템의 세부 연구 내용

이와 같이 본 연구에서는 사용자가 간단한 조작으로 영상을 재생성하여, 손쉽게 비디오 콘텐츠를 제작할 수 있도록 하는, 영상 저작 도구를 구현하는 것을 목표로 한다. 제안 시스템은 크게 3가지 모듈로서, 영상 분류 처리기, 이미지 검색 추출기, 그리고 합성 영상 생성기로 구성된다. 특히 합성 영상 생성기는 합성 대상에 따라 객체, 얼굴, 배경의 3가지를 구분하여 처리하도록 하였으며, 객체 합성의 경우, 적응적 트라이맵을 사용한 새로운 합성 알고리즘을 제안한다. 이와 같이 본 논문의 연구내용에 대한 서술은 주로 각 모듈을 구현하기 위한 기법들에 대해서 논의하며, 시스템 설계와 구현 과정을 설명한다. 그리고 새롭게 제안한 기법을 소개하고, 아울러 성능과 실험 결과를 통해 고찰한다.

3. 논문구성

이 논문은 다음과 같이 구성된다. 2 장에서는 제안 시스템 개발에 필요한 관련 연구에 대해 논의하고, 3 장에서는 전체 제안 시스템에 대해서 상세히 소개한다. 그리고, 4장에서는 새롭게 제안하는 적응적 트라이맵을 이용한 합성 영상 기법에 대해서 자세히 설명한다. 기존 상용 시스템과의 비교 결과는 5장, 결론은 6장에서 기술한다.

II 관련 연구

본 장에서는 본 연구에서 구현한 적응적 트라이맵을 사용한 합성 영상 편집 시스템을 소개하기에 앞서, 이와 관련한 기존 기술들을 기술한다. 첫째, 사용자 참여 기반의 콘텐츠 제작 기술들을 살펴보고, 둘째, 제안 시스템의 영상 분류처리기 개발과 관련한 영상 분류처리기술에 논의한다. 그리고 셋째, 이미지 검색 추출기 개발과 관련된 기반 기술을 정리하고, 마지막으로 넷째, 객체, 얼굴, 배경을 나눠서 영상을 합성하기 위한 합성 영상 생성기 개발과 연관된 기존의 방법들을 살펴본다.

1. 사용자 참여 기반의 콘텐츠 제작 기술

사용자 참여 기반의 콘텐츠 제작 기술 개발에 있어서, 산업계와 대학연구실 및 국책연구기관은 저차원(텍스트,이미지,오디오) 콘텐츠는 기술 안정화에 진입한 상용 서비스가 제공되고 있지만, 고차원(3차원, 동영상) 기술은 상용화 이전의 원천기술 개발에 집중적으로 연구가 활발히 이루어지고 있다. 주로 Social TV, 모바일 SNS와 증강 현실 기반의 새로운 플랫폼에서 응용하기 위한 실험들이 진행되고 있다. 대표적인 사용자 참여 기반의 콘텐츠 제작 기술과 관련된 실험적인 연구는 다음과 같다.

1.1. Social TV 서비스 플랫폼

한국전자통신연구원(ETRI)에서는 차세대 IPTV 인프라 구축에 힘입어, IPTV와 Social(소셜) 서비스의 장점을 결합한 소셜TV 서비스 제공을 위한 인에이블러를 개발하기 위한 연구를 진행하였다[1]. 이 연구

를 통해 콘텐츠 개발자 뿐만 아니라 서비스 사용자들도 효율적이고 자유롭게 소셜TV 플랫폼을 통하여 콘텐츠를 저작하고, 유통할 수 있게 하는 서비스 인에이블러 기술을 개발하였다. 대표적인 소셜TV 서비스 인에이블러 기술 개발 내용으로는 도메인,주제,키워드 사이의 의미를 모델링하는 알고리즘을 도출하는 소셜 정보 기반 사용자 관심사 구조화 기술 연구를 시작으로, 그림6과 같이 이미 구축되어진 콘텐츠 메타데이터의 모델링 데이터와 사용자 프로필과 같은 사용자 모델링 데이터, 그리고 영상에서 표정으로 추출할 수 있는 사용자 감정정보 등 이러한 3가지 모델을 결합하여 TV 콘텐츠 연관도를 모델링하는 기술 제안하였다. 이를 응용하여 사용자들에게 특화된 콘텐츠를 추천할 수 있도록 하는 기능을 개발하였다.

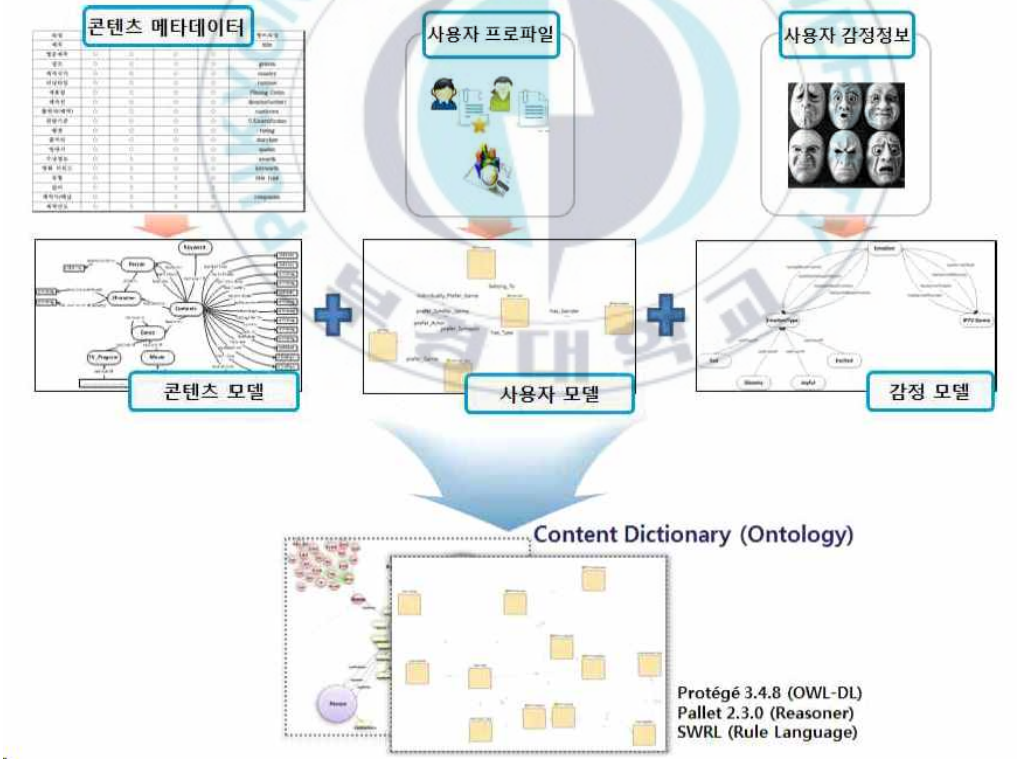


그림 6. 콘텐츠 사전을 이용한 사용자 연관 맞춤형 영상 제공 기법

또한 연관키워드를 이용한 협업 filtering 추천 알고리즘은 TF-IDF(Term Frequency-Inverse Document Frequency) 값 기반의 CF(Collaborative Filtering)를 활용하여 데이터을 추출하고, 추출한 데이터를 이용하여 연관 콘텐츠를 검색 및 추천할 수 있도록 하여, 서비스 목적에 부합하는 추천결과를 도출하고자 하였다. 또한 이러한 연구결과물은 기존의 IPTV 셋탑과 연동 소셜 TV 서비스를 안드로이드 플랫폼에 상용 서비스를 할 수 있도록 Seceond-Screen서비스의 형태로도 개발하였다.

이러한 시도는 갈수록 눈높이가 높아지는 사용자 요구에 부합하기 위해서, 사용자들이 원하는 시나리오에 따라 맞춤형 영상을 제작할 수 있다는 점에서 큰 의미를 가진다. 또한 이와 같은 기술은 콘텐츠 검색이나 추천에 응용할 수 있다는 점에서 확장성을 가지는 장점도 있다. 반면에, 의미와 내용을 포함하는 메터 데이터를 만들고, 태깅하는 기술의 난해함으로 인해 종종 사용자가 원하는 영상 콘텐츠가 아닌 오류를 나타내기도 한다. 하지만, 기본의 자료와 정보를 활용하여, 사용자가 손쉽게 직접 영상을 재생성하고 편집할 수 있게 된다면, 사용자가 원하는 맞춤형 영상을 제작에 있어 가장 만족도가 높을 것이라 예상되어진다.

1.2. 사용자 참여형 증강현실(AR) 콘텐츠 제작 플랫폼

서울시립대 황지은 교수 연구실에서는 사용자 참여형 모바일 증강현실 콘텐츠 제작하고 활용하기 위한 필요 요소와 실증적 한계를 도출하였다[2]. 특히 사용자 경험을 통해 보완할 수 있는 다양한 방법을 연구함으로서 기초제계를 마련하고자 하였다. 대표적인 연구로서, 그림 7과 같이 원격현장(remote-site)에서는 데스크탑을 활용한 웹기반 콘텐

츠 제작 시스템을, 반면에 현장(on-site)에서는 모바일 기반 콘텐츠 체험 시스템으로 구성하여, 두 현장에서 순환되는 증강현실 콘텐츠의 제작 체계를 정의하고, 이를 바탕으로 사용자 경험(UX)을 위한 실험을 수행하였다.

그림 7. 사용자 참여형 모바일 증강 현실 콘텐츠 제작 시스템

위 그림 7과 같이 사용자 참여기반의 증강현실(AR) 콘텐츠 제작 플랫폼은 크게 2가지로 구성되며, 하나는 데스크탑 기반의 원격현장(remote-site)웹서비스이고, 다른 하나는 모바일 기반 현장(on-site) 애플리케이션이다. 데스크탑 기반 웹서비스에서는 증강현실 대상을 트래킹하기 위해서, 위치 정보나 마커 정보를 활용해 인터액티브하게 시나리오를 입력할 수 있게 하였다. 이렇게 제작된 콘텐츠를 가상현실에서 실제 바로 사용가능하도록 개발하였다. 현장(on-site) 애플리케이션

은 다양한 콘텐츠 정보를 실제 대상에 증강현실처럼 체험할 수 있도록 하였고, 아울러 현장 정보를 기반하여 콘텐츠를 재구성하였다.

이와 같은 실험은 사용자 참여 가능하도록 위치와 마커 정보를 활용하여 증강 현실 컨텐츠를 접목시킨점에서 아주 훌륭한 시도임에는 틀림이 없지만, 이미지와 텍스트 정보를 적용하여 콘텐츠를 제작하는 것에 비해 새로운 영상 정보를 제작하는 것에는 한계가 있음을 보여준다.

1.3. Social TV와 SNS 영상 공유를 위한 스트리밍 기술

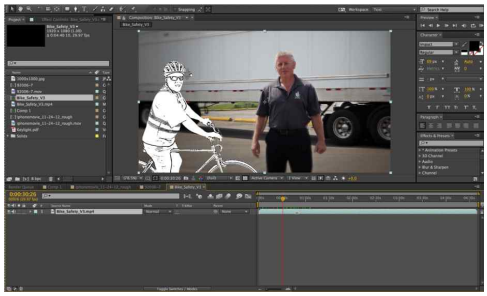
광운대학교 정광수 교수 연구팀에서는 사용자 맞춤형 영상의 수요 증대에 따른 트랙픽 문제를 해결하고, 영상을 효율적으로 전송하는 기법을 적용한 시스템을 제안하였다[3]. 이는 다양한 단말 및 네트워크 환경에서 동작하는 특징을 고려하여, 모든 사용자들의 서비스 품질을 보장하기 어렵다는 단점을 해결하고, 서비스의 품질을 향상시키기 위한 적응적인 스트리밍 기법을 제안하였다. 이 시스템은 소셜 그룹 내의 사용자들이 사용하는 단말 특성과 사용자가 위치한 네트워크 특성을 인지하여 적응적으로 콘텐츠 품질을 조절하도록 하였다. 또한 MPEG-UD를 사용하여 수집한 시청자 컨텍스트를 구조화하고, 이러한 정보를 바탕으로 사용자 맞춤형 시나리오를 제공하여, 결과적으로 사용자에게 맞춤형 영상을 제공할 수 있는 컨텍스트 처리 시스템을 제안하였다[4]. 이는 시청자의 반응이나 의도에 따라 영상 미디어의 상호작용 서비스 제공할 수 있음을 보여주었다.

2. 영상 분류 처리기 개발 관련 기술

본 연구의 제안 시스템을 구성하는 영상 분류 처리기는, 기존 상용 영상 편집기를 응용 발전시켜, 사용자들의 참여 가능한 새로운 저작도구로 구현하는 것이 목표이다. 이에 먼저 개발되어 서비스 중인 상용 시스템을 분석하고, 아울러 기존 시스템에 적용되는 기법과 공개된 연구 결과틀 토대로 살펴본다. 이 장은 다음 3가지 사항에 중점을 두고 서술한다. 첫째, 영상 합성 편집과 단순 사용자 조작에 특화된 시스템을 구성하기 위해서 사용자 인터페이스에 중점을 두고 분석하며, 둘째, 영상 합성의 대상 구간을 설정하는 기능을 구현하기 위해서 장면 전환 기법에 대해서 논의한다. 마지막으로, 셋째, 이미지 합성 시 객체의 위치를 결정하기 위해 적용 가능한 관련 객체 트랙킹 기법을 살펴본다.

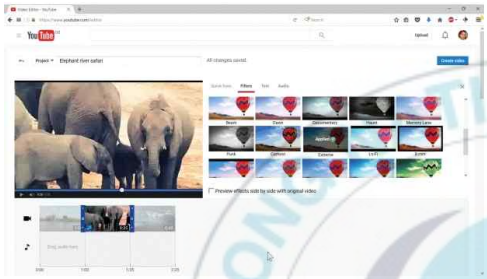
2.1. 비디오 영상 분류와 편집을 위한 제작 도구

기존의 상용 영상 편집 도구들은 그림8과 같이 대상 사용자가 .전문가이냐 일반인이냐에 따라 크게 2가지로 분류할 수 있다. 각각 2가지씩 대표적인 도구를 예를 들어보면, 하나는 숙련된 조작 기술을 보유한 전문가들이 사용하는 아래 그림8-1.Abobe의 After Effect와 그림8-2.Sony의 Vegas 가 대표적이고, 다른 하나는 간단한 조작법만 익히더라도 일반인 누구나 손쉽게 사용할 수 있는 그림8-3.YouTube의 Video Editor와 그림8-4.VideoToolbox의 Movavi 가 대표적이다.



1. Adobe의 After Effect

2. Sony의 Vegas



3. Youtube의 Video Editor

4. Video Toolbox의 Movavi

그림 8. 기존의 상용 영상 편집 시스템

이 두 가지 분류 중 먼저 전문가 조작 도구를 살펴보면, 1.After Effect는 타임라인 위에 레이어를 사용하여 편집하는 방식으로서 영상 합성에 우수한 품질과 성능을 보이며, 모션 그래픽스와 같은 고급 기능들도 처리를 한다. 또한 각 객체는 개별 트랙을 가지는 레이어 지향 방식이므로 영상의 확장과 키프레임 처리에 강점을 보인다. 반면 2.Vegas는 레이어가 아닌 트랙 지향 의 비선형 편집도구로서, 하나의 영상 객체는 한 트랙을 점유해서 사용하기 때문에, 영상을 잘라내고, 붙이고 하는 등의 컷 편집 기능에 강점을 보인다. 이 두 가지 도구 중 본 연구의 제안시스템과 같이 영상 합성을 주목적으로 하는 편집도구의 개발에서는 After Effect와 같은 레이어 기반 방식이 적합하다고 판단되어진다.

또한 1,2 이 2가지 편집도구는 공통적으로 전문가의 숙련된 조작 능력이 필요하므로, 복잡한 사용자 인터페이스와 메뉴 도구들로 구성되어 있다. 이에 본 논문의 제안 시스템은 합성과 대상 구간 설정과 같은 핵심 기능으로 구성된 단순한 사용자 인터페이스를 구현하여 사용자 조작을 손쉽게 하고자 하였다.

다음으로 사용자에 따른 2가지 분류 중, 일반인들이 사용하는 편집 도구로서 첫 번째 3.Youtube Editor는 웹 기반 어플리케이션으로서 영상을 다듬고, 잘라내고, 여러 영상을 결합하는 등의 편집을 할 수 있다. 그리고 두 번째 4. Movavi 도 마찬가지로 합치기, 자르기, 꾸미기와 필터 효과와 같은 작업을 할 수 있는 특징을 가진다. 그리고 이 2가지 도구는 공통적으로 웹을 통한 배포가 쉽다는 특징과 무엇보다 조작이 간편해서 사용이 쉽다는 장점을 가진다. 하지만, 영상 재생성을 위한 합성 기능은 제공되지 않는 단점이 있다. 제안 시스템에서는 이들 시스템과 같이 조작 편의성을 살리고, 또한 웹을 통한 연동, 배포와 같이 영상과 이미지를 저장하고 관리할 수 있도록, 구현 시스템 모듈의 이미지 검색 추출기와 연동하도록 기능을 구현하였다.

2.2. 비디오 영상의 자동화된 장면 전환 검출 기술

장면 전환 알고리즘은 비디오 영상을 분할하기 위해 영상을 구성하는 최소 단위인 샷을 찾아내는 기술로서, 프레임에서 특징을 추출하는 방법을 정의하여 전환 지점을 찾아내거나, 클러스트화된 메타 데이터나 자막과 같은 특정 정보를 기반으로 비교하거나, 프레임간의 유사도를 측정하여 전환 지점의 샷을 검출하는 방법등이 있다[5-7]. 영상 관련 기술의 성능을 측정하고 검증하는 TRECVID에서는 샷검출은 물리적 특

성의 변화만으로도 손쉽게 처리할 수 있는 기법들이 발표되었다[8].

그림 9. 장면 전환 검출을 사용한 영상 편집 시스템 예시

이와 같이 본 논문의 제안 시스템의 영상 분류 처리기에서는 합성 편집 대상 구간을 장면 단위로 구성하여 작업하도록 하고, 편집 시작 위치를 설정하면, 장면 검출 기법을 통해 마지막 위치를 자동으로 설정하도록 기능을 구현하였다.

2.3. 합성 위치 선정을 위한 모션 트랙킹

영상 기반의 모션 트랙킹은 영상의 추적 시작 프레임에서의 객체부터 움직임을 추정해서 연속되는 이후 프레임에서 같은 객체를 찾아내는 과정이다. 가장 대표적인 알고리즘은 다음과 같다. 첫째, 가우시안 잡음이 포함된 선형 시스템에서 원 신호를 예측하는 칼만 필터[9], 둘째, 다수의 파티클 샘플로 예측값을 만들어, 이 예측값 샘플에서 관측값 확률을 구해서 추정값을 계산하는 몬테카를로 방식의 파티클 필터[10], 그리고 이전프레임과 다음프레임의 한 픽셀의 밝기의 시간에 따른 변화량, 즉 방향과 이동거리, 속도로서 움직임에 대한 플로우를 사용하여 추정하는 옵티컬 플로우가 있다[11,12]. 그리고 객체의 위치정보가 아닌 초기 지정 영역의컬러 히스토그램을 이용한 색 구성의 확률 분포를 계산해 탐색윈도우를 위치시켜 유추하는 평균이동 알고리즘인 Meanshift[13,14]와 여기에 탐색윈도우의 자동 크기 조절을 기법을 추가해 개선시킨 Camshift 알고리즘[15,16]이 있다. 이러한 대표적인 객체 추적 알고리즘을 위시하여 각각의 알고리즘을 상황이나 목적에 따라 결합하고 개선시켜 연구되었다. 예를 들어 배경영상 복잡도, 초기 입력정보의 유무와 같은 상황이나 실시간성, 얼굴 추적등의 목적에 따라 그에 적합한 기법들이 제안되었다. 본 연구의 제안시스템에서는 모션 트랙킹의 목적은 신규 객체의 합성 처리 후, 다음 프레임에서의 기존 객체의 위치를 추적하여, 신규 객체의 위치를 결정하기 위함이다. 다시 말해, 다음 프레임에서 추정한 기존 객체의 위치는 신규 객체의 합성 대상 위치를 기준 위치를 의미한다. 본 논문에서는 이와같이 모션 트랙킹 기법에 관한 연구는, 해당 기능, 즉 제안 시스템에서의 합성 대상 위치 기능 구현을 목적으로 한다.

3. 이미지 검색 추출기 개발 관련 기술

본 논문에서 개발하는 제안 시스템의 이미지 검색 추출기는 합성 영상을 재생성하기 위한 전처리 단계로서, 합성 가공되지 않는 원본 영상을 찾거나, 혹은 원본 영상에 합성할 대상 이미지나, 영상을 검색한 후 원하는 이미지를 추출하는 작업을 한다. 이 절에서는 이와 관련된 기존의 기반 기술들을 살펴본다.

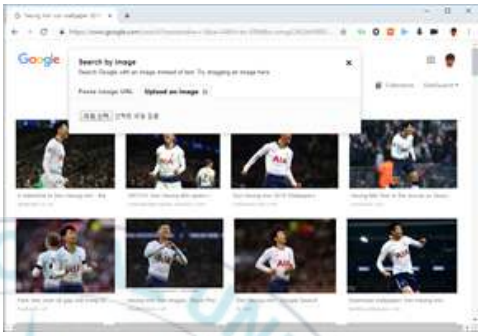
3.1. 이미지 검색 기술

이미지 검색 기술은 파일이나 데이터베이스에 저장된 이미지들 중에서 자신이 원하는 이미지를 찾아내는 기술이다. 이 기술은 크게 2가지 범주로 구분할 수 있다. 첫 번째는, 텍스트로 구성된 속성정보인 메타데이터나 태깅 정보를 이용하여 이미지를 찾아내는 텍스트 기반 이미지 검색 방식이고[17], 두 번째는, 이미지를 구성하는 색상(color), 텍스처(texture), 형태(shape) 등의 특징을 이용하여 이미지를 검색하는 내용 기반 이미지 검색 방식[18]이다.

텍스트 기반 이미지 검색은 해당 이미지와 추가적으로 입력된 제목(title), 내용(context), 그리고 해쉬태그 등의 텍스트 정보를 연관시켜 검색하기 때문에, 입력 과정에서 연관정보를 알맞게 입력만 한다며, 정확한 검색을 보장 할 수 있다. 또한 텍스트 검색은 검색 속도가, 이미지 기반 검색에 비해 빠르다. 반면에 이미지와 상관없는 텍스트가 타이틀이나 태그로 저장되어 있으면 검색 정확도가 떨어지는 단점도 존재한다. 이러한 문제는, 결국 저장된 값에 의존하기 때문에, 사용자가 텍스트를 직접 입력하는 방식에서는 입력 후 검색 연관성이 낮은 텍스트를 제거하고, 수정보완해서 처리하고, 반면 텍스트가 자동 생성되어 태

깅 되는 방식에서는 입력단계부터 정확한 텍스트를 입력되도록 처리된다. 또한 이미지에 저장된 속성 정보, 즉 메터 데이터를 이용하기도 하는데, 주로 이미지의 촬영 날짜, GPS 좌표로 표현되는 위치 정보를 사용한다.

1.입력 이미지



2.검색 결과

그림 10. 내용 기반 검색 방식의 예

내용 기반의 이미지 검색 방식은 그림10과 같이 사용자가 이미지를 입력하면, 그와 유사한 이미지를 검색 결과로 출력해 주는 방식으로서, 태그과 제목과 같은 사용자의 추가 텍스트 입력을 필요로 하지 않는다. 대신 이미지 자체정보인 색상(color), 텍스처(texture), 형태(shape)와 같은 특징을 사용하기 때문에, 검색의 정확도는 이러한 특징들의 검출 정확성과 매칭 알고리즘 성능에 비례한다. 색상 기반 검색 방식[19]은 누적 색상 히스토그램이나 도미넌트 색상 특징을 검출하여 검색하는 방식이고, 텍스처 기반 검색 방식[20]은 이미지 내에서 많은 부분을 차지하는 영역에 반복적으로 나타는 패턴의 질감을 특성으로 이용하는 방식으로서, 주로 2가지 방법으로 텍스처 특징을 검출한다. 첫 번째는 웨이블릿과 같은 주파수 분석으로 특징을 검출하는 방법이고, 두 번째는 GLCM(Gray Level Cooccurence Matrix), LBP(Local Binary Patterns)

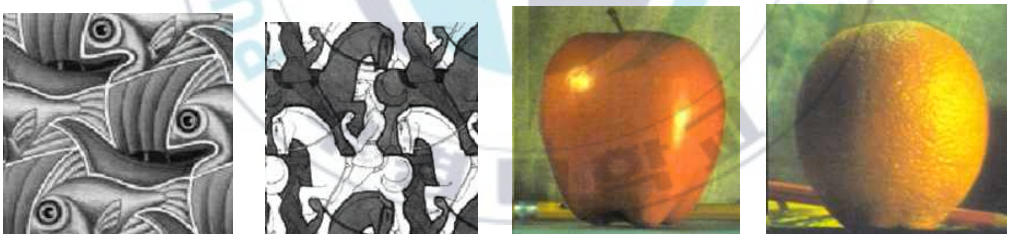
와 같이 통계 분석 으로 특징을 검출하는 방법이다. 형태 기반 방식[21]은 이미지 내의 객체와 의미 영역의 폐쇄된 경계 영역을 특징으로 하여 검색하는 방식으로서 크게 2가지 방식으로 구분된다. 첫 번째는, 푸리에변환이나 히스토그램을 사용하는 컨투어(Contour) 기반 방식이고, 두 번째는, Zernike 모먼트나 Gometry 모먼트와 같은 영역(Region) 기반 방식이다.

4. 합성 영상 생성기 개발 관련 기술

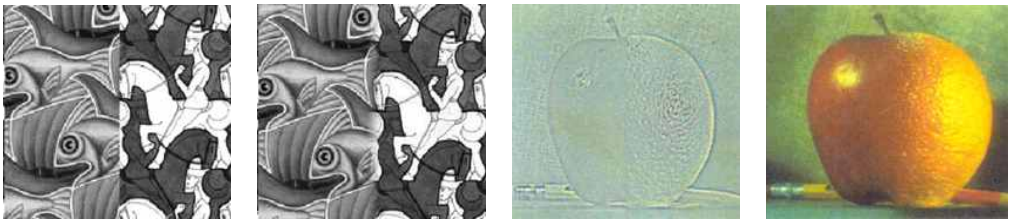
제안 시스템에서 구현된 합성 영상 생성기는 전체 시스템 모듈 중 가장 핵심 역할을 하는 처리기로서, 원본 영상을 가공하여 새로운 영상으로 재성성하기 위해, 영상을 구성하는 프레임 이미지를 합성하는 연산을 한다. 또한 제안 합성 처리기는 합성 대상의 종류에 따라 구분하여 처리하는데, 이를 통해 합성 품질과 연산속도, 그리고 사용자 조작의 편의성을 향상시키고자 한다. 합성 대상은 객체, 얼굴, 배경으로 3가지로 구분한다. 이 절에서는 이와 관련된 기존의 기반 기술들을 살펴본다.

4.1. 객체 합성 기법

이미지 합성(compositing)은 크게 이미지 혼합(blending)과 이미지 매팅(matting) 2가지로 범주로 나눌 수 있다. 먼저 이미지 혼합은 그림 11의 예와 같이 좌우 2개의 이미지를 붙여서 합성하는 것이다.



1-1.이미지1 1-2.이미지2 2-1.이미지1 2-2.이미지2



1-3.좌우 단순붙임 1-4.좌우 합성 2-3.라플라시안 2-4.좌우 합성

그림 11. 좌우 이미지 혼합(blending)

그림 11. 1처럼 경계선 부분의 필터 적용에 퓨리에 변환을 이용하여 알파값을 결정하여 매끄럽게 처리하거나, 그림11. 2와 같이 라플라시안 피라미드를 적용시켜 혼합(blending)한 이미지를 만들어 낸다.

또 다른 예로 그림12처럼 같은 장소의 다른 특징, 여름 겨울과 같은 계절의 차이, 낮과 밤, 날씨와 같이 차이나는 영상을 서로 혼합하기도 한다.

1-1.영상1(여름) 1-2.영상2(겨울) 2-1.이미지1(어둠) 2-2.이미지2(밝음)

1-3.상하 합성1 1-4.상하 합성2 2-1.알파 값 조정 2-4.상하 합성

그림 12. 상하 이미지 혼합(blending) - 1.계절(날씨), 2.시간(밝기)

그림 12의 1-3은 위쪽은 눈이 덮이 겨울 영상2 이고, 아래쪽은 물이 보이는 여름 영상1 이다. 두 가지를 상하의 위치에 혼합한 것을 볼 수 있다. 그림 12의 1-4는 반대로 혼합한 것을 볼 수있다. 그림12의 2-4는 위쪽은 2-1 이미지1을, 아래쪽은 2-2 이미지2를 알파값 조정을 통해 혼합한 결과이다.

이미지 합성에 있어서 이미지 혼합(blending)은 주로 2가지 이미지를 좌우, 상하로 합치거나 장소가 비슷한 곳의 다른 특징을 섞어서 새

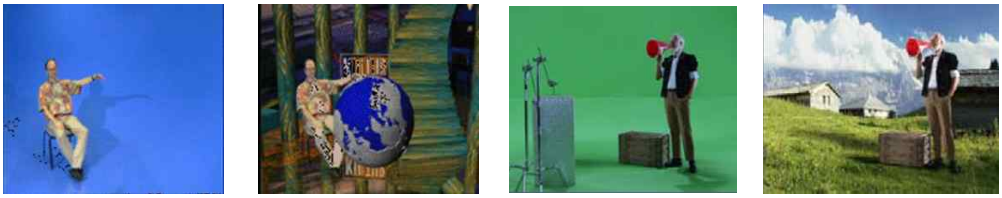
로운 이미지를 만드는 목적으로 쓰인다. 그래서 주로 경계선의 자연스럽게하기 위한 연구가 많이 진행되었다[23].

이미지 합성 기법 중 2번째 이미지 매팅(matting) 기법은 그림 13과 같이 하나의 이미지에서 대상 객체를 추출해서, 새로운 배경의 이미지 위에 위치시키고 합성 시키는 것이다.

1. 원본 이미지 2. 전경 객체 추출 3. 새로운 배경에 합성

그림 13. 이미지 매팅의 예

이미지 매팅의 시초는 그림 14와 같이 배경을 그린 혹은 블루스크린으로 놓고 합성시키는 크로마키 기법이다. 이는 고전적인 방법이지만, 아직도 방송이나 영화에서 많이 사용되고 있다. 하지만 배경이 반드시 단순한 색상으로 구성되어 있어야 하기 때문에, 이미 만들어진 복잡한 영상을 합성 편집하는데는 유용하지 않다. 그래서 이후 복잡한 배경에서 이미지를 추출하여 합성시키는 많은 매팅 기법들이 제안되었다[22].



1-1.블루 스크린 1-2.합성결과 2-1.그린 스크린 2-2.합성결과

그림 14. 크로마키 기법

이미지 매팅 기법의 대부분은 그림 15와 같이 입력 이미지와 사용자가 입력한 트라이맵(Tri-map) 이 2 가지를 가지고, 알파매트(alpha matte)를 구해내는 일이다. 트라이맵이란 그림15-2에서 보듯이 확실한 배경(Backgoround)과, 확실한 전경(Foreground), 그리고 알 수 없는(Unknown)의 3가지 영역을 구분한 맵이고, 알파매트는 이 트라이맵의 알 수 없는(Unknown)영역의 알파값을 결정한 후 생성된 매트를 의미한다. 결국 이 알파매트를 정확하게 구하는 작업이 대부분의 이미지 매팅의 목표이다.

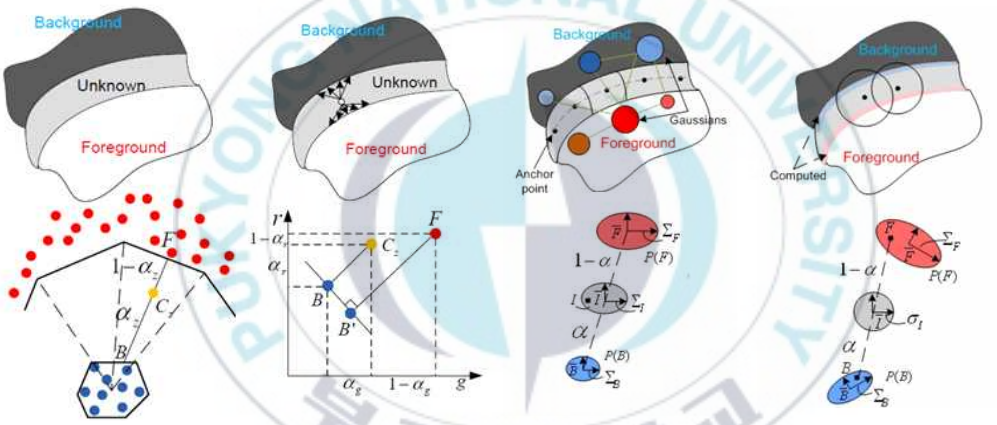
1. 입력 이미지 2.트라이맵(Tri-map) 3.알파매트(alpha_matte)

그림 15. 일반적인 이미지 매팅 기법

이미지 매팅 기법은 크게 3가지 범주로 나눌 수 있는데, 첫 번째는, 색상 샘플링 기반(color sample based)방식이고, 두 번째는 어피니티 기반(affinity-based)방식이고, 그리고 마지막으로 세 번째는 색상 샘플링 기반 방식과 어피니티 기반 방식을 결합한 혼합(combined) 방식이다.

색상 샘플링 기반 방식은 그림 16과 같이 대표적으로 4가지가 있는데, 첫째, 기본적인 블루 스크린 적용 모델로서 단순히 전경(F)와 배경(B)의 색상을 구분하여 알파값(α)을 적용하는 색상(C)를 구해내는 그림16-1.Mishima[23]가 제안한 기법이 있고, 둘째, 전경과 배경의 경

계선에 인접한 픽셀 거리에 비례하는 가중치 합(weight sum)의 평균으로 색상 표본(sample)을 구해, 알 수 없는(Unknown) 영역의 알파값(α)을 추정하는 그림16-2.Knockout[24]방식이 있다. 그리고 세 번째로, Ruzon과 Tomasi가 제안한 방법으로 알 수 없는(Unknown)영역의 각 객체의 색상 분포(distribution)의 시작점들을 프로티어(frontiers)로 정의하고 이와 연결된 영역을 통로(manifold)를 따라 알파값을 측정하는 그림16-3.Ruzon-Tomasi 방법[25]이 있다. 이 때 색 표본들은 무방향 가우시안(non-orientated Gaussians)에 의해 모델되어진다.



1.Mishima[23]	2.Knockout[24]	3.Ruzon,Tomasi[25]	4.Bayesian[26]
블루스크린 기하 모델	내추럴 매팅 기하 모델	내추럴 매팅 확률 모델	내추럴 매팅 베이지안 모델

그림 16. 색상 샘플링 기반 매팅 기법

마지막으로 네 번째, 그림16-4. 베이지안(Bayesian) 매팅 기법[26]은 트라이맵의 인접한 픽셀에 의존한 확률 모델로 알 수 없는(Unknown)의 알파값을 찾아 경계선 문제를 푸는 방법이다. 이는 세 번째 Ruzon-Tomasi 방법과도 유사하며 색 표본들이 방향성을 가지는 가

우시안 확률 분포로 추정된다는 점에서는 차이가 난다. 또한 베이지안 매팅은 매트를 구할 때 최대 사후 확률(maximum a posteriori)을 사용한다.

이와 같이 색상 샘플링 기반의 매팅 기법들은 복잡한 장면 속의 색상 샘플링은 분류 오류(mis-classification)가 나타나는 본질적으로 문제가 존재한다. 그래서 이러한 문제를 해결하기 위해서 이웃 픽셀들 사이의 다양한 조화특성(affinities), 예를 들어 매트 기울기(matte gradient)를 정의하여 이미지의 지역적 통계를 적용하는 기법이 제안되었다.

어피니티 기반(affinity-based)방식은 대표적으로 그림17과 같은 포아송(Poisson) 매팅 기법[27]과 그림 18과 같은 클로즈폼(Closed Form) 매팅 기법[28]이 있다. 먼저 그림 17과 같은 포아송(Poisson) 매팅은 전경과 배경의 밝기값의 변화를 매트 기울기(gradient)로 모델링 하고, 포아송(Poisson) 방정식을 적용하여 매트 추출 오류를 줄이는 방법이다. 하지만 이 방식은 밝기 값의 변화, 즉 기울기가 평활하다는 전제하에서만 우수한 품질을 보장한다.

1.벡터 필드 가이던스(v)　　2.그라디언트 필드(g)　　3.Unknown(Ω)보간

그림 17. 포아송(Poisson) 이미지 매팅

클로즈 폼(Closed-form) 매팅은 위에서 설명한 포아송(Poisson) 매팅에서 전경과 배경에서 밝기값의 기울기가 지역적으로 평활하다는 전제하에 유도된 비용함수의 문제점을 해결하기 위해서, 알파값(α)값의 방정식에 희소 선형 시스템을 적용하여 해결하고자 하였다. 그림18-2와 같이 사용자가 이는 기존의 사용자 입력인 트라이맵과 달리, 사용자가 붓터치입력(Scribbles)을 통해 매트를 추출하였다. 또한 밝기값의 기울기가 아니라 이미지 윈도우(image window)를 사용한다는 점에서 포아송(Poisson) 기법과 차이를 가진다. 하지만 이 기법 또한 전경과 배경은 지역적으로 선형이라는 조건에서만 우수한 성능을 보장한다.

1.입력 이미지 2.사용자 입력 3.매트 추출 4.객체추출 5.배경 합성

그림 18. 클로즈 폼(Closed from) 매팅

색상 샘플링 기반 방식과 어피니티 기반 방식의 문제점을 해결하기 위해서 이 두 가지를 결합한 혼합(Combined)방식으로 Robust 매팅 기법[30] 이 있다. 이 기법은 클로즈 폼(Closed from) 매팅과 컬러 샘플링 기법을 결합한 것으로서, Random Walk 방법[30]을 사용한 매팅 에너지 함수의 최소화로 색상 샘플의 신뢰값(confidence)을 향상시켰다.

그림19은 텍스처를 생성하는 그래프컷(graph cut) 알고리즘[31]으로, 이후 디지털 이미지와 비디오 영상을 합성하기 위한 목적으로, 사용자의 인터랙티브한 입력 조작을 통해, 전경을 추출할 수 있는 그랩컷(Grabcut)방법[34]에 응용하여 적용된다.

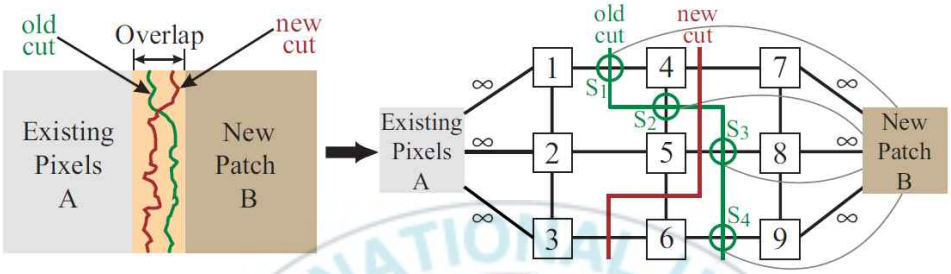


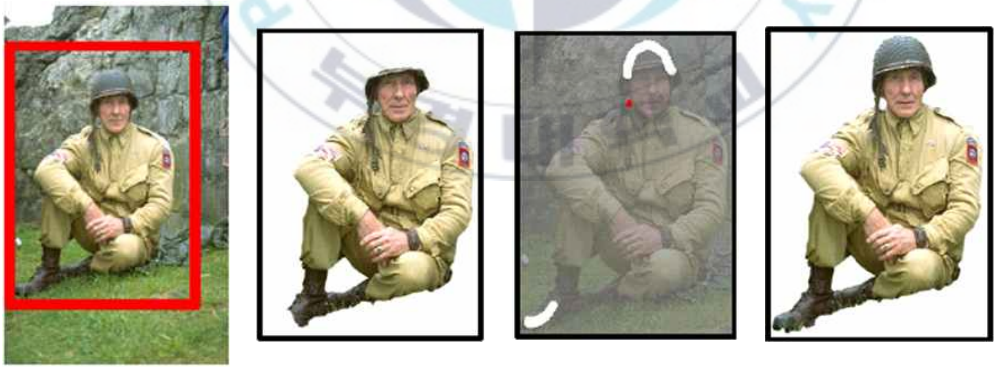
그림 19. Graphcut 알고리즘

이전의 샘플기반(sample-based)합성 기법은 해당 샘플과 시각적으로 유사한 영상을 새롭게 생성하는 기법이고, 패치기반(patch-based) 합성 기법은 샘플 영상의 일부분을 단 번에 복제하는 채워넣는 기법이다. 하지만 이 두 가지 기법의 공통적인 복제된 이미지 패치들의 이음새가 끊어지는 문제점을 나타냈다. 이러한 문제점을 극복하기 위해서, 그래프컷 알고리즘에서는, 패치사이의 경계를 찾아내서, 시각적으로 끊김이 없는 새 이미지를 생성하는 기법을 제시하였다.

그리고 그림 20과 같이 Interactive digital photomontage[33]에서는 그래프 컷[31]과 포아송(Poisson) 이미지 매팅[32]을 결합하여 합성하는 방법을 제안하였다.

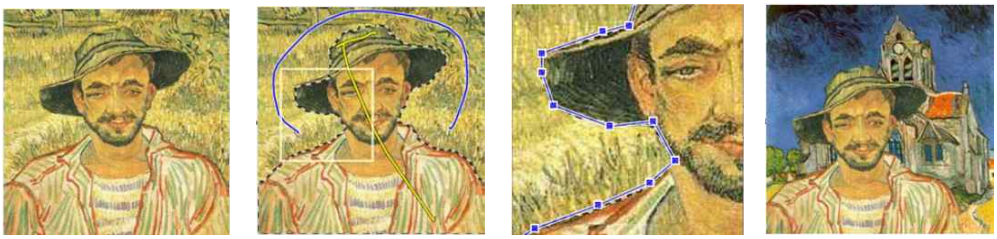
그림 20. Interactive Digital Photomontage

그 외 매팅 기법을 사용자 인터페이스 관점에서 살펴보면 트라이맵을 사용하지 않는 방법은 그림 21과 같이 Grab Cut 방법[34]와 그림 22와 같은 Lazy Snapping 기법[35]이 있다. 이 두 기법은 모두 그래프 컷(graph cut) 알고리즘을 사용하는 공통적인 특징이 있지만, 그랩컷(Grab Cut) 기법은 배경에만 사각 영역선택 윈도우나 라쏘(Lasso)영역 선택 도구를 사용하고, 반면에 Lazy Snapping 기법은 전경과 배경 모두에 마킹 브러쉬(marking brush)를 사용할 수 있고, 추가적으로 버텍스(vertext)를 조정이 가능하다는 차이가 있다.



1.자동 사각형 2.객체 1차 추출 3.사용자 마킹 4.객체 최종 추출

그림 21. Grab Cut 기법



1.입력이미지 2.객체 마킹 3.경계선 조정 4.합성 결과

그림 22. Lazy Snapping 기법

지금까지 논의한 객체 합성을 위한 이미지 매팅 기법들을 제안 시스템에 적용하기 위해서 크게 3가지 사항에 중점을 두고 연구하였다. 첫 번째는, 합성 품질의 결과를 높이는 방법, 두 번째는, 사용자 개입의 편의성을 위한 사용자 인터페이스, 그리고 마지막 세 번째, 합성 영상 재생성 편집 시스템에 적합한 수준의 자동화이다. 본 논문에서는 이와 같은 사항들을 고려하여 시스템을 구현하고자 하였다.

4.2. 얼굴 합성 기법

제안 시스템의 얼굴 합성은 원본 영상의 얼굴을 새로운 이미지의 얼굴로 교체(swap)하는 기능을 의미한다. 따라서, 선행적으로 원본 영상과 교체대상 이미지에서의 얼굴 검출하고 특징점 추출하는 작업이 중요하다. 현재 얼굴 검출(Face Detection) 기술은 전통적인 연구자 직접 설계 특징(hand-crafted features) 기반 방식에서 시작하여, 딥러닝(deep learning) 기반 방식으로 진화하였다.

먼저 연구자 직접 설계 특징(hand-crafted features) 기반 방식은 추적(Casecade) 방법[36], DPM(Deformable Parts Model) 기법[37], 그리고 ACF(Aggregated Channel Features) 알고리즘[38]과 같이 크게

3가지 범주에 포함된다. 가장 근본적으로 Haar-like 특징을 사용하는 AdaBoost 기술로 작동하는 Viola-Jones 얼굴 검출기[39]가 대표적이며, 이 검출기와 유사한 구조에 기반하여 새로운 SURF[40], HoG[41], 그리고 LBP[42]와 같은 새로운 특징들을 추가시킨 검출기들이 있다. 이와 같이 대부분의 얼굴 검출기들은 연구자가 직접 설계한 특징(hand-crafted features) 여러 가지를 조합하는 방식으로 정확성을 높이고자 하였다. 이러한 전통적인 방식의 얼굴 검출 기법은 연산속도가 빨라서 대부분 실시간성을 보장하지만, 복잡한 얼굴 변화(자세, 표정, 조명)에서는 강인하지 못하다는 단점이 있다. 그래서 품질이 나쁜 저해상도의 사진이나 영상에는 적합하지 않다고 평가된다.

그래서 이러한 문제를 극복하기 위해서, 수많은 이미지들에서 더 다양한 변화 요소들을 학습시킨 딥러닝(Deep learning) 기반의 방식들이 제안되었다. 이러한 연구들은 특히 아주 작은 얼굴까지도 추출하려는 시도를 하고 있다.

딥러닝에 기반한 얼굴 검출 기법은 학습 알고리즘에 따라 CNN(cascaded convolutional networks)[43], faster R-CNN[44], 그리고 SSD(Single shot multibox detector)[45]와 같이 크게 3가지 범주로 나눌 수 있다. 이 방식들은 전통적인 방식에 비해 상대적으로 우수한 결과를 보장하지만, 연산 속도가 느리다는 단점이 있다. 이는 검출 성능을 높이려고 할수록, 연산 시간이 오래걸리는 트레이드오프(tradeoff)가 존재한다.

그림23은 얼굴 검출과 베이스라인 생성 조정 기법을 결합한 예를 보여준다.

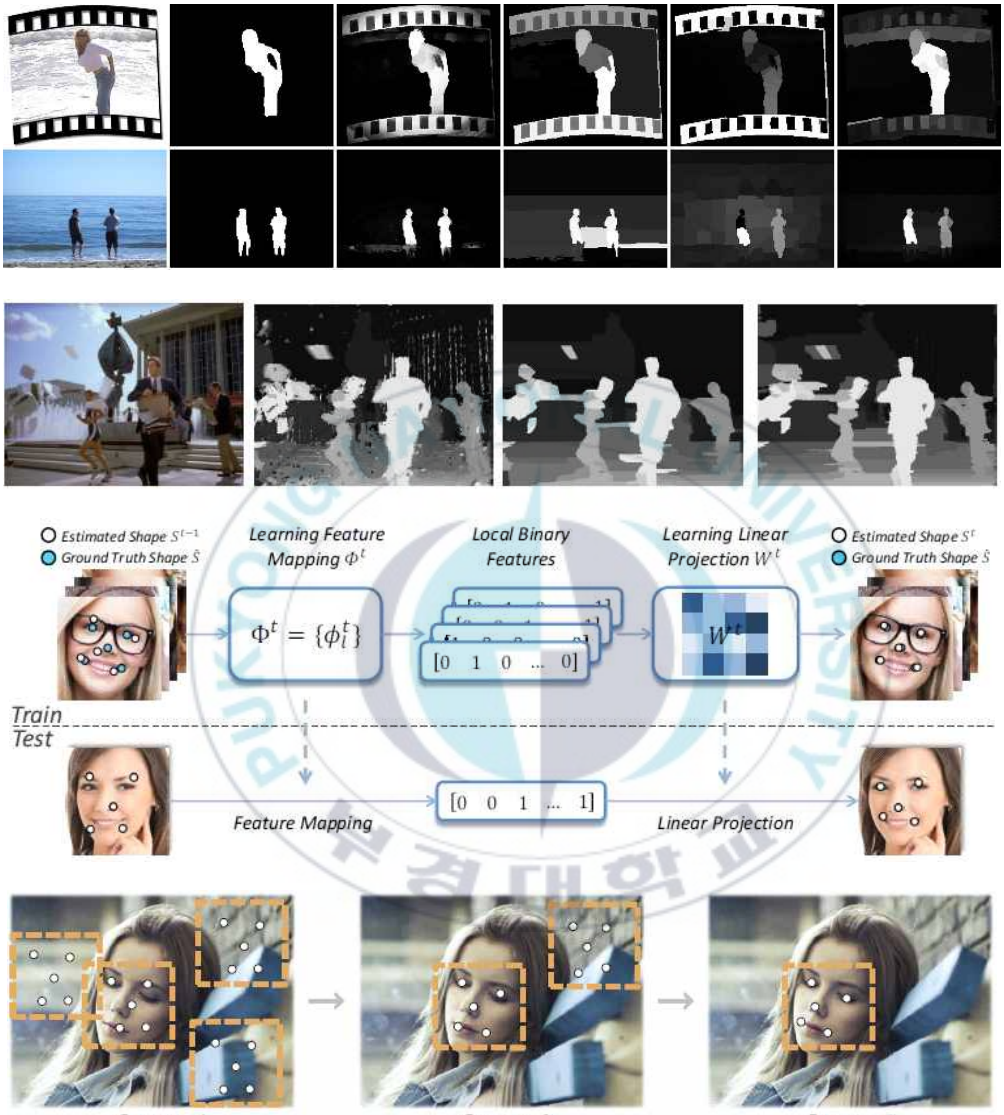


그림 23. salient 얼굴 검출을 위한 베이스 라인 기법과 얼굴 검출

4.3. 배경 합성 기법

제안 시스템의 배경 합성을 간단하 표현하면 원본 배경을 떼내고, 새로운 배경으로 교체하는 작업이다. 배경을 떼어내기 위해서는 이미지로부터 객체와 배경 영역을 분리하는 작업이 선행되어야 한다. 다음 그림 24는 전통적인 기존 연구들에서 영역을 분리하기 위한 기법들을 구분한 것이다. 먼저 첫 번째, 1.Classification은 이미지에서 관심 객체(예:고양이)의 포함여부만을 단순히 판단하는 기법이고, 두 번째, 2.Localization은 객체를 둘러싸는 경계사각(bounding box)과 객체 위치까지 구하는 기법이다. 그리고, 세 번째, 3. Detection은 해당 객체들을 경계사각 수준으로 위치를 모두 검출하는 것을 의미한다. 마지막으로 네 번째, 4.Segmentaion 은 그림24-4와 같이 객체를 둘러싼 정확한 경계까지 구하는 기법이다.

1.Classification 2.Localization 3.Detection 4.Segmentation

그림 24. 이미지 영역 분리 기법

본 논문의 배경 영상 합성에서는 위의 기법 중 4.Segmentation 기법을 통해 객체 영역을 구분한 후, 객체를 제외한 배경 영역에 새로운 배경을 입힌다. 따라서, 영상의 정확한 분석을 통해, 이 구성요소들 주요 성분들을 분해하고, 해당 물체의 위치와 외곽선을 찾아내는 영역 분할 작업이 중요하다. 이러한 영역 분할(Segmentation) 기법은 픽셀(pixel), 에지(edge), 영역(region)에 기반 방법에 따라 3가지로 나눌 수 있다. 첫 번째, 픽셀 기반 방식은 히스토그램을 이용 픽셀 분포에 따

라 임계값(thresold)을 설정하여 경계선을 구해내는 기법이고, 두 번째, 에지(Edge) 기반 방식은 에지 검출 필터를 사용하여 경계선을 추출한 후, NMS(Non-Maximum Suppression)로 의미없는 에지를 제거하는 식으로 구현된다. 마지막으로 영역 기반 방식은 영역 구성의 동질성(homogeneity)을 정의한 후, 의미있는 영역으로 나누고자 하는 것인데, 어떻게 동질성을 규정하는가의 문제가 관건이다. 이러한 방식은 대개 영역 확장(region growing) 방법으로, 시드(seed)라 불리는 픽셀 초기 시작점에서 점점 영역을 늘려가는 방식을 취한다. 처리과정의 따라 그림 25와 같이 상향식(Bottom-up) 방식과 하향식(Top-down) 방식로 나뉜다. Bottom-up 방식은 유사한 특징 요소들을 묶는 방식이고, Top-down 방식은 동일 객체들을 집단화 하는 것이다[46].

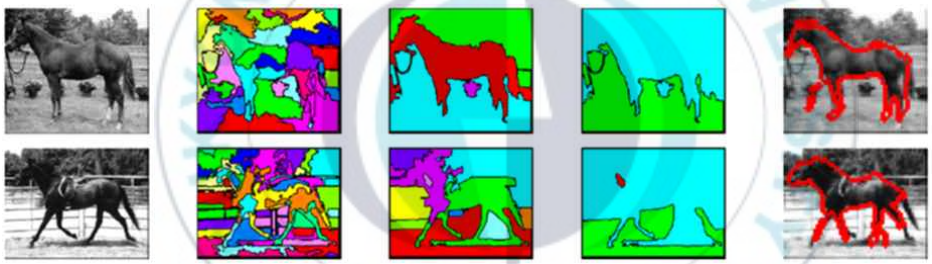


그림 25. Bottom-up, Top-down Segmentation

외외에도 영역 단위로 합병(merge), 나누기(split) 처리를 조합한 방식으로 구현된다. 본 연구에서는 이와 같은 영역 분할 기법들을 분석하여서, 배경 합성 시에 필요한 교체배경을 추출하기 위해서, 적합한 방법을 선정하고, 응용한 후 구현하였다. 제안 시스템은 영역분할 기법과 사용자의 최소한의 개입이 필요한 사용자 인터페이스와 결합하여 최적의 배경 분할 결과를 얻고자 하였다.

III 제안 시스템

1. 시스템 개요

본 논문에서 제안하는 영상 재생성 시스템은 기본적으로 콘텐츠 제공자들에 의해 만들어진 원본 소스 영상과 추가적으로 태그 기반 검색을 통하여 얻어진 후보 영상에서 추출된 이미지를 기반으로 하여, 사용자 처리과정을 거쳐 도출된 영상과 이미지를 보정하는 합성 처리를 통하여 최종적으로 새로운 영상을 생성한다. 제안 기법은 크게 3단계의 처리과정을 거친다.

첫번째는 원본 영상을 소스로 하여 변경할 개체를 추출하는 1단계 영상 분류 처리기이고, 두번째는 합성 대상이 되는 영상을 검색하고, 그 영상에서 필요한 이미지를 추출하는 2단계 이미지 검색 추출기이고, 마지막 세번째는, 프레임 비교와 보정을 통하여 최종 수정된 합성영상을 생성해내는 합성 영상 생성기이다. 이와 같이 3단계 처리기로 구성된다.

1단계의 영상 분류 처리기에서는 원본 소스 영상을 가지고, 유사 이미지들로 구성된 프레임들의 참조 구간을 설정한다. 장면 전환을 기준으로 하여, 변경 구간들의 후보를 인덱스 한다. 또한 변경을 위한 배경 범위나 관심 객체들로 구성된 구간들을 선택 할 수 있도록 구성한다. 참조 구간의 해당 프레임 마다 변경 영역의 개체를 추출한다. 이 경우 유사하지만 다양한 형태의 개체들이 추출되어지게 되는데, 예를 들어, 얼굴의 경우, 다양한 각도와 표정으로 나타나게 된다. 이러한 여러 장의 이미지를 조합하여 3차원 모델로 변환하는 작업을 수행한다.

2단계의 이미지 검색 추출기에서는 사용자 입력 텍스트와 후보 영상 저장소에 저장되어진 다양한 영상들의 태그를 서로 비교하여 1차적으로 후보 영상들을 뽑아낸다. 이후 영상에서 사용자가 원하는 프레임

을 선택한 후, 원본 소스 영상에 합성하기 위한 대상 이미지 개체를 추출 한다. 이는 단순한 사각형 형태의 배경 이미지가 될 수도 있고, 복잡한 다각형 형태를 가지는 얼굴, 객체(예:비행기)와 같은 이미지가 될 수도 있다.

마지막 3단계의 합성 영상 생성기에서는 2단계에서 추출된 대상 이미지를 1단계에서 설정한 구간에서의 프레임과 하나씩 비교한 후, 1단계 추출 개체와 이미지 변환을 통하여 최종적으로 수정된 합성 영상을 생성한다. 이때 얼굴합성의 경우 1단계에서 변환 생성된 3차원 모델을 가지고, 사용자가 원하는 적절한 포즈의 따른 2차원 이미지를 만들어 냄으로서 새로운 형태의 영상을 재생성 할 수 있다. 또한 추가적으로 합성 부분의 경계선의 일그러짐을 보정하여 자연스러운 영상을 만들어 낸다.

사용 사례 설명을 위해서, 기존 원본 영상의 스토리 전개를 '기본 시나리오' 라고 정의를 하고, 사용자에 의해 변경되는 내용을 '사용자 시나리오' 라고 정의를 한다. 사용자는 기존 원본 영상에 자신이 참여한 시나리오를 적용하는데 이는 데이터베이스로 구축된 저장소에서 텍스트 기반으로 하여 후보 영상을 검색하고, 영상 중에서 사용자 시나리오를 적용하여 변경할 이미지를 추출하게 된다. 결국 추출된 이미지와 기본 영상의 합성은 사용자 시나리오를 기본 시나리오에 적용하여 적용된, 다시 말해 수정된 시나리오가 적용된 결과 영상이 생성된다. 이는 사용자가 직접 영상 재생성에 참여하여 상호 반응형 맞춤형 콘텐츠를 공급할 수 있는 환경을 제공한다. 또한 여러 장의 이미지를 모델화하는 시도를 통하여 새로운 영상을 재생성 함으로서, 사용자들이 보다 더 다양한 경험을 할 수 있는 기회를 제공하고자 하였다.

본 연구에서 구현한 3단계 핵심 개발 모듈을 포함한 전체 시스템 구성은 아래 그림 26과 같다.

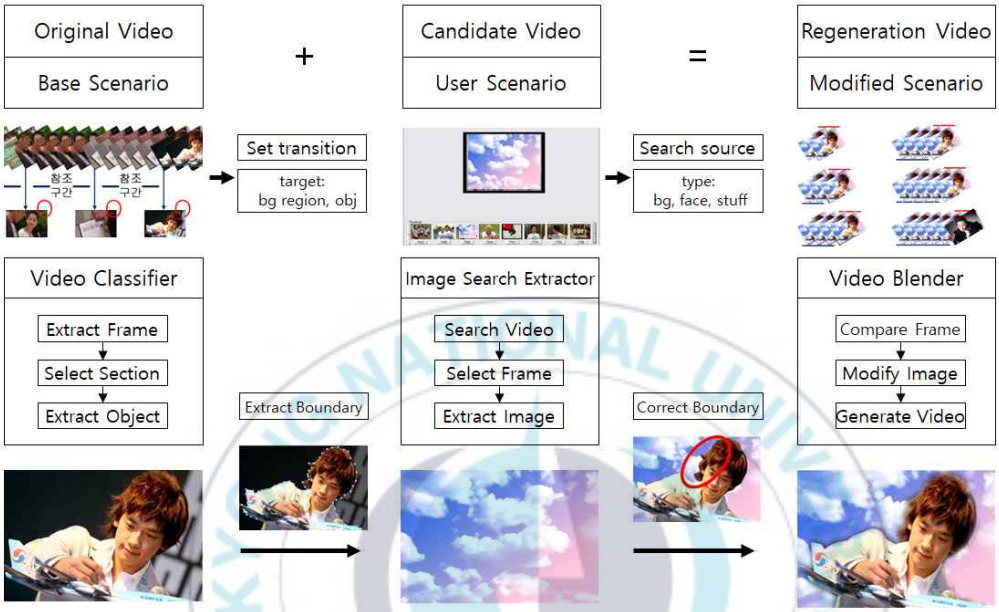


그림 26. 전체 결과 시스템 구성도

2. 영상 분류 처리기

영상 분류 처리기는 기본 시나리오로 정의된 원본 소스 영상에서 유사한 이미지들로 구성된 편집 대상 참조 구간을 설정하고, 변경하고자 할 프레임을 추출한 후, 이미지 검색 추출기를 통해 선정된 이미지를 합성 하는 일련의 과정들을 처리하는 통합 편집 저작도구로서 아래 그림 27과 같이 구현하였다.

그림 27. 제안 시스템의 영상 분류 처리기

2.1. 사용자 인터페이스

제안 시스템의 사용자 인터페이스(User Interface)는 비숙련자도 최소한의 설명만 듣고, 단순 조작만으로도 손쉽게 다룰 수 있도록 직관적인 조작 도구를 구현하는데 초점을 맞춰 개발하였으며, 전체 조작 도구와 그 주요 기능은 아래 표2로 정리하였다.

전체 조작 도구 구성은 조작 컴포넌트(버튼, 슬라이더바, 콘텐츠 작업뷰)의 배치에 의해 크게 3영역으로 구분된다. 첫번째, 그림27의 좌측에 위치한 작업도구들은 영상과 이미지의 로딩과 선택을 위한 콘텐츠 관리 도구와 콘텐츠 재생성 작업대상 모드 즉, 객체, 배경, 얼굴등을 선택하는 모드 지정 도구, 그리고 최종 합성을 처리하는 합성 처리 도구로 구성된다. 두 번째, 우측 상단부를 차지하는 영상과 이미지 뷰는 원본 영상과 교체될 이미지를 동시에 보여주면서 작업 할 수 있도록 디스플레이 도구와 마우스 영역 설정을 처리하는 이벤트 처리도구, 그리고 바로 하단에 작업 진행 중인 영상 확인을 위한 재생 도구로 구성된다.

마지막으로, 세 번째, 우측 하단부는 영상과 이미지를 타임라인 기반의 레이어로 편집할 수 있도록 하는 레이어 관리 도구와 콘텐츠 구간 설정 도구로 구성된다.

	기능 분류	기능 조작 컴포넌트	설명
1	콘텐츠 선택 및 로딩	Select Video	원본 비디오 영상 선택
2		Select Image	교체될 이미지 선택
3		Select Template	클립보드의 콘텐츠 선택
4		Load Video	영상 파일 불러오기
5		Load Image	이미지 파일 불러오기
6	변경 대상 구간 설정	Add Movie(+)	영상을 편집창에 추가
7		Add Layer(+)	이미지를 편집창에 추가
8		Subtract Movie(-)	영상을 편집창에 제거
9		Subtract Layer(-)	이미지를 편집창에 제거
10		Layer Control Slider	대상 구간 설정
11		DEL	대상 구간 삭제
12	영상 합성 처리	Lasso Tool	객체 추출 도구 모드
13		Cut	객체 추출 작업 완료
14		Move Object	객체 이동
15		Blending	객체 합성 처리
16		Select Ground	배경 합성 설정 모드
17		Select Target	얼굴 합성 설정 모드
18	합성 영상 재생	Play	합성 영상 재생
19		Stop	합성 영상 재생 종료
20		Pause	합성 영상 일시 정지
21		Record	합성 영상 녹화

[표 2] 구현 시스템 사용자 인터페이스 설명

위의 표2와 같이 기능 조작 컴포넌트는 기능에 따라 4가지 정도의 그룹으로 분류할 수 있다. 첫 번째로, 콘텐츠 선택 및 로딩 그룹에서는 주로 편집 대상이 되는 영상과 이미지들을 파일에서 불러와서 프로젝트 창에서 관리 할 수 있도록 하는 콘텐츠 관리도구로서, 탭뷰를 사용하여 구분하였고, 썸네일 아이콘으로 표시되는 탐색기 형태로 보여줌으로서 파일을 직관적으로 선택하고 제거할 수 있도록 구현하였다. 또한 영상과 이미지들을 다중으로 로드하여 준비시켜 놓음으로서 편집시에 일괄 작업이 가능하도록 편의성을 제공하고자 하였다. 두 번째, 변경 대상 구간 설정 그룹에서는 주로 (+)버튼을 이용하여 새로운 레이어를 생성하거나, (−)버튼을 이용하여 제거하는 기능을 수행할 수 있다. 또한 레이어의 우선순위를 설정하고 표시하는 기능들로 구성하였다. 우선순위는 레이어의 기본적인 상하 배치 순에 따라서 결정이 되며, 임의로 텍스트 에디터에 인덱스 값을 입력함으로서 조절할 수도 있다. 그 외에, 각각의 레이어들은 개별적인 슬라이더바로 생성되어, 각 영상과 이미지들의 디스플레이될 시작과 끝을 손쉽게 설정할 수 있도록 하였다. 세 번째, 영상 합성 처리 기능 그룹에서는 본 시스템의 가장 중요한 역할인 객체, 배경, 얼굴 합성 작업을 처리 할 수 있다. 이 기능을 위해서는 우선적으로 대상 모드를 선택하는 작업이 선행되어야 하면, 이는 각 Select 버튼으로 구현하였다. 객체 합성 처리 모드의 경우 Lasso Tool 버튼으로, 배경 합성 처리 모드의 경우 Select Background 버튼으로, 그리고 마지막으로 얼굴 합성 처리 모드의 경우는 Select Target 버튼으로 설정할 수 있다. 객체 합성의 경우 교체할 이미지로부터 영역을 추출하는 도구가 필요한데, 직관적인 사용을 위해서, 그림 27의 우측 상당 이미지 디스플레이 뷰에서 마우스 클릭으로 직접 선택 한 후, 최종적으로 Cut 버튼을 통하여 영역 선택을 완료할 수 있다. 그리고, 이어서 move버튼을

통해서 원본영상에 추출한 이미지의 위치를 설정하고, 동시에 마우스 휠로 이미지의 크기를 조정할 수 있도록 하였다. 그 외 배경 합성 처리의 경우도 교체할 배경을 선택할 시에 객체 합성시 방법과 마찬가지로 이미지 디스플레이 창에서 마우스 클릭으로 선을 그어서 설정할 수 있도록 하였다. 마지막으로, 합성 영상 재생 그룹은 편집 진행 중에 상태를 점검할 수 있도록 재생, 종료, 정지 기능을 버튼으로 구현하였고, 완료된 영상을 저장할 수 있도록 녹화 기능를 구현하였다.

2.2. 편집 대상 구간 자동 설정

구현 영상 분류 처리기에서 합성 편집 대상 구간은 장면 단위로 구성하여 작업할 수 있도록 하였다. 이 경우 각 레이어별로 존재하는 사용자 인터페이스의 슬라이더바를 통해 편집구간의 시작 위치를 설정하면, 장면 검출 기능을 통해 마지막 위치를 자동으로 추천하고, 사용자 확인 후 최종 설정할 수 있도록 구현하였다.

1. 편집 시작 지점 설정　　2. 장면 전환 구간 자동 검출

그림 28. 합성 편집 대상 구간 설정을 위한 장면 전환 지점 검출

구현 시스템의 장면 전환 기능은 기본적으로 HSV 색상 공간에서 두 프레임간의 변화에 기반한 내용기반(content-aware) 검출 방식과 추가적으로 프레임 intensity/brightness 평균에 기반한 임계값 입력 방식으로 구현하였다.

2.3. 영상 합성 시 객체 위치 선정

제안 시스템의 영상 처리 분류기에서는 원본 영상의 한 프레임에 사용자가 원하는 객체를 합성 후, 이후 프레임에도 객체를 움직이는 객체의 추적을 위해서 CAMshift와 Kalman Filter를 결합하는 방식으로 구현하였다.

비디오 영상에서 선택 객체를 변환하여 재성성하기 위해서는, 현재의 프레임(이미지)에서 다음 프레임(이미지)의 객체의 위치를 추적해야 필요가 있다. 다음 순간의 위치를 추정할 때, 보통 선택된 객체의 위치와 속도 등의 정보를 바탕으로 값을 구한다. 대표적인 CAMshift 알고리즘의 경우, 선택 객체의 영역, 즉 탐색 윈도우를 설정할 때, 이전 프레임에서의 객체의 위치 정보를 바탕으로 사각형의 가로와 세로 크기를 확대해서 구한다. 하지만 이 경우 고속 이동객체의 경우 추적의 정확성이 떨어지는 문제가 있다. 그 외에 Kalman filter 알고리즘의 경우는 순환적 선형 구조로 되어 있는 필터로서 수식을 통해 미래정보를 예측해 낸다. 이때 예측 과정에서 현재 프레임의 위치 측정값이 항상 필요하다는 특징이 있다. 이는 이 값을 구하지 못하면 추적을 못하게 되는 문제가 있다.

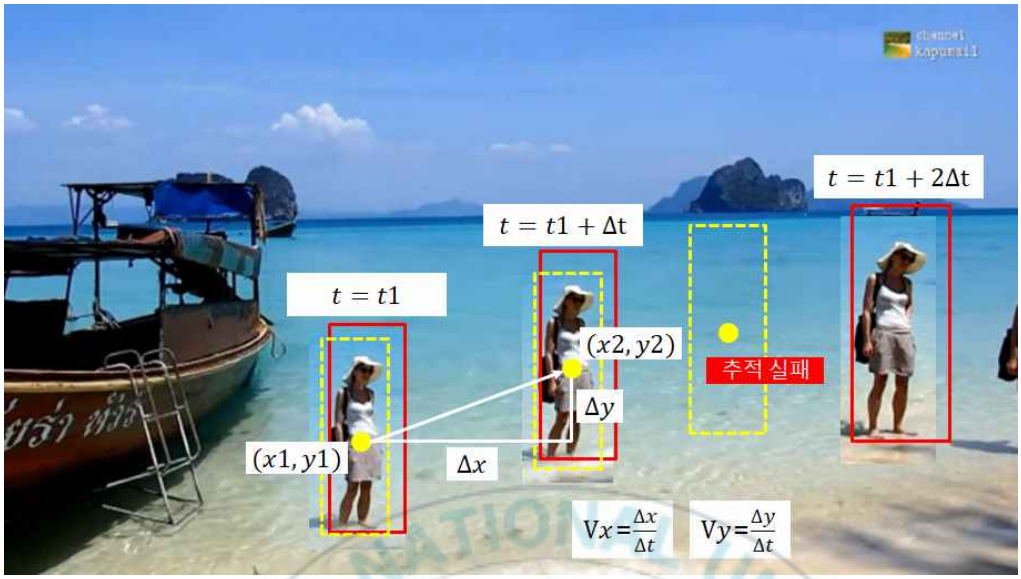


그림 29. 추적 실패의 예와 추적 요소 값들

그림 29 는 고속 이동 객체의 추적 실패를 표현한 것이며, 노란색 점선 사각형은 CAMshift 탐색윈도우를 나타내며, 빨간색 사각형은 Kalman 탐색 윈도우를 나타낸다.

제안 시스템에서는 이러한 문제를 보완하기 위해서, 두 알고리즘 (CAMshift와 Kalman filter)을 결합한 방법으로 객체를 추적한다.

$$\hat{x}_k = [x,\ y,\ \Delta x,\ \Delta y\ ]^{\mathrm{T}} \quad \text{.....(1)}$$

$$y_k = [x,\ y\ ]^{\mathrm{T}} \quad \text{.....(2)}$$

$$u_k = [v_x\ ,\ v_y\ ]^{\mathrm{T}} \quad \text{.....(3)}$$

두 알고리즘의 결합을 위해서 시스템에 맞게 상태벡터, 측정벡터, 제어벡터와 같은 요소를 새롭게 모델링 하였다. 위 식 (1)과 같이 상태벡터는 물체의 좌표, 즉 중심벡터의 변화량을 의미한다. 식 (2)와 같은

측정 벡터는 Camshift값에 수렴하는 측정된 물체의 좌표를 나타낸다. 식(3)과 같이 제어벡터는 현재 속도 정보로 구성된다.

이와 같은 객체의 추적을 위해서 요소를 모델링하여, CAMshift와 Kalman Filter를 결합하는 방식으로 구현하였다.

그림 30은 CAMshift 와 칼만 필터를 결합한 추적 알고리즘의 처리 과정을 나타낸다.

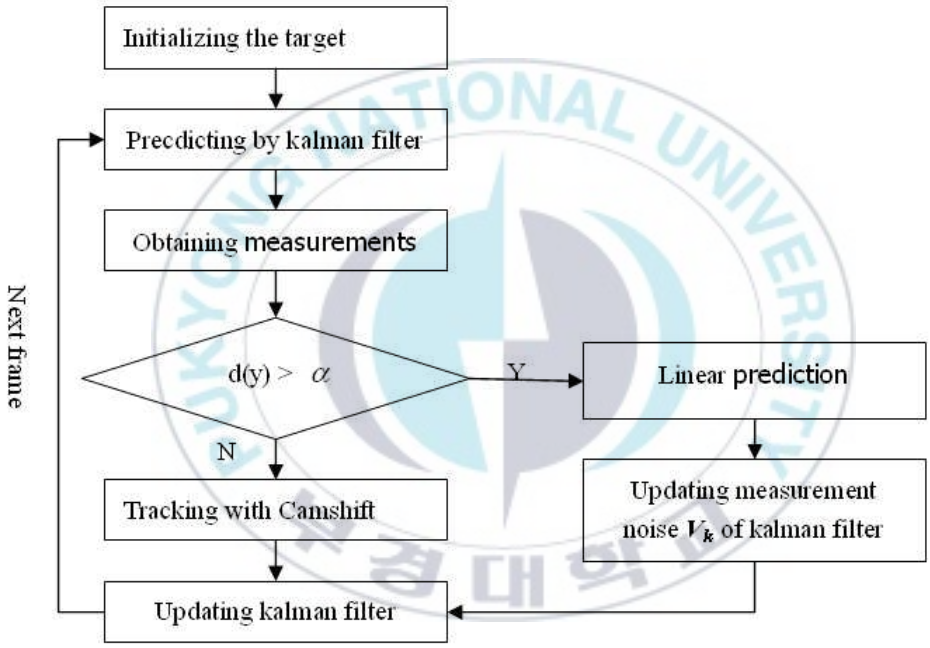


그림 30. 움직임 객체 추적 알고리즘 순서도

아래 그림 31은 구현 시스템의 객체 추적 결과를 통해 위치를 선정한 후의 합성 결과물을 나타낸다.

그림 31. 객체 추적 결과를 통한 합성 대상 위치 선정

3. 이미지 검색 추출기

구현 시스템의 이미지 검색 추출기는 사용자가 태그와 같은 입력 텍스트로 데이터베이스화 되어 구축된 저장소 플랫폼에서 다양한 이미지와 영상을 검색한 후 추출하는 작업을 한다. 영상의 경우 사용자가 원하는 프레임을 선택한 후, 원본 소스 영상에 합성하기 위한 대상 이미지를 추출한다. 제안 시스템에서 구현한 이미지 검색 추출기는 그림 32와 같다.

그림 32. 제안 시스템의 이미지 검색 추출기

제안 이미지 검색 추출기는 텍스트 기반 검색 방식으로 구현하였으며, 저장 이미지에 해당하는 제목, 내용, 메타정보와 같은 기본 속성을 저장하는 기본 테이블과 다중 해쉬태그를 저장하는 별도의 릴레이션 테이블로 구성한 관계형 데이터베이스 시스템을 구축하였다. 이미지와 영상은 그림 33과 같이 별도의 파일서버로 구성된 저장소에서 관리하도록

하였으며, 파일명을 기본테이블에 넣어 관리하도록 하였다. 그 외에 유튜브나 비메오 같은 상용 서비스의 영상도 검색할 수 있도록, 영상의 URL들을 기본 테이블의 컬럼으로 구성하여 관리하도록 하였다.

그림 33. 제안 시스템의 웹 기반 이미지 및 동영상 저장 플랫폼

4. 합성 영상 생성기

제안 시스템의 합성 영상 생성기는 그림 34와 같이 합성 대상에 따라 객체, 얼굴, 배경으로 구분하여 합성 작업을 하도록 하였다.

그림 34. 대상물 분류에 따른 객체 합성 결과

4.1. 객체 합성

제안 시스템에서 정의하는 객체는 이미지를 구성의 3가지 분류 중에서 얼굴, 배경과 구분지어, 이들을 제외한 나머지 대상 물체를 의미한다. 객체 합성 방법은 단순한 인터페이스로 구성되며, 사용자 합성 조작 방법은 다음과 같이 2가지 순서로 작업하도록 구성하였다. 첫 번째로 라쏘선택도구(Lasso tool)로 소스 영상에서 사용자 객체의 영역을 따낸다. 그리고 두 번째, 원본 영상에 마우스 드래그(Drag)로 추출 객체를 합성 위치로 이동시키고 및 와 마우스 휠 신축(Scaling) 조작으로 객체의 크기 조절한다.

제안 시스템에서는 객체 대상 합성 연산시 포아송(Poisson) 매팅 기법[32]을 적용하여, 경계선(boundary)을 매끄럽고 자연스럽게 처리하고자 하였다. 이 알고리즘의 접근법은 2가지로 나누어진다. 첫째, 입력 이미지와 사용자가 선택한 트라이맵(trimap)의 경계선 정보를 이용하여 뽀아송 근사된(approximated) 매트의 밝기값의 기울기 필드(gradient field)를 구해낸다. 둘째, 이렇게 얻어진 밝기값의 기울기 필드를 뽀아송 방정식으로 연산해서 최종 매트를 구해 낸다.

먼저 첫째, $\nabla\alpha$ 근사로 밝기값의 기울기 필드 구한다.

식(1)의 I 는 합성되어지는 새로운 이미지를 의미하고, F 는 전경 이미지, 그리고 B 는 배경 이미지를 의미한다. 식(1)의 양변을 편미분하여 근사 매트의 밝기값의 기울기 필드(gradeint field)를 구해낸다.

$$I = \alpha F + (1-\alpha)B \quad \cdots\cdots(1)$$

$$\nabla I = (f-B)\nabla\alpha + \alpha\nabla F + (1-\alpha)\nabla B \quad \cdots\cdots(2)$$

여기서 전경 F 과 배경 B 의 각각의 R,G,B 채널을 더해서 채널수로 나눈 밝기값 변화가 평활한 경우, 식(3)의 $\alpha\nabla F + (1-\alpha)\nabla B$ 는 $(F-B)\nabla\alpha$ 이 비교적 작기 때문에 근사 매트의 그레디언트 필드를 식(4)와 같이 구할 수 있다.

$$\alpha\nabla F + (1-\alpha)\nabla B \approx 0 \quad \cdots\cdots(3)$$

$$\nabla\alpha \approx \frac{1}{F-B}\nabla I \quad \cdots\cdots(4)$$

이것은 매트의 그레디언트가 이미지의 그레디언트에 비례하다는 것을 의미한다. 이와같은 매트 그레디언트의 근사기법은 Mitsunaga[49]

에서 처음 고안되었다. 솔리드 객체의 경계선 주변의 투명도를 추정하기 위해서, 솔리드 객체의 경계선에 수직인 1차원 경로를 따라서 매트의 그레디언트를 적분하였다. 뽀아송 이미지 매팅에서도 동일한 근사 방법을 사용하지만, 매트를 더 효율적으로 재구성하기 위해서 2차원 이미지의 평면상에서는 직접적으로 뽀아송 방정식을 풀어낸다.

두 번째, 뽀아송 방정식을 이용하여 최종 매트 구한다.

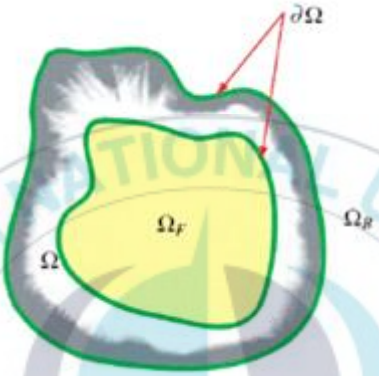


그림 35. Poisson 매팅

그림35와 같이 Ω_F는 사용자 선택한 분명한 전경이고, Ω_B는 분명한 배경이고, Ω는 구하고자 하는 영역(unknown region)으로 정의된다. 이 이미지에서 각 픽셀 $p=(x,y)$에 대해서, I_p는 그 픽셀의 밝기를 나타내고, F_p는 전경의 밝기, 그리고, B_p 는 배경의 밝기를 의미한다. N_p는 주변의 4개 픽셀이며, $\partial\Omega=\{p\in\Omega_f\cup\Omega B|N_p\cap\Omega\neq 0\}$는 Ω 의 외부경계를 의미한다.

주어진 근사 매트 $(F-B)$와 이미지의 그라디언트 ∇I를 이용해 Ω 영역의 매트를 구하기 위해서, 다음의 변화량 문제의 최소화값을 구한다.

$$\triangle \alpha = div \left(\frac{\nabla I}{F - B} \right) \quad \text{.....(5)}$$

$$\triangle = \left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} \right) \quad \text{.....(6)}$$

동일한 경계 조건을 가질 때의 뽀아송 방정식은 식 (5)와 같이 정의되며, 여기서 div 는 발산연산자(Divergence operators)를 의미한다. 식(6) 라플라시안(Laplacian)이다. 제안 시스템에서 뽀아송 방정식은 논문[28]에서도 사용된 가속완화법이 포함된 가우스-자이델(Gauss-Seidel) 반복법을 사용한다. 이 때 컬러 이미지들은, 그레이 스케이 채널에서 $(F-B)$와 ∇I 둘 다 측정되어진다.

1. 합성 결과2 (overlap)　　2.합성 결과3 (방법3)

그림 36. 포아송(Poisson) 기반 객체 합성 과정

뽀아송 방정식을 푸는 것에 의해 이미지를 합성하는 뽀아송 매팅은, 사용자의 반자동(semi-automatically)입력과 근사화된 매트와 그라디언트 필드를 측정하여 매트를 재구성하는 기법이다. 이미지 합성시 샘플링 오차로 인해 발생하는 문제를 줄이고, 빠르고 품질 좋게 결과를 얻어낸다. 이는 또한 국부 지역의 연산을 통해 몇 가지 힌트만으로도, 이전의 기법들로 풀지 못했던 복잡한 이미지 합성에 대해서 그림 37과 같은 인상적인 결과물을 얻을 수 있다. 하지만 밝기값이 평활하다는 전제하에만 품질을 보장하기 때문에, 복잡한 이미지에서도 합성 결과를 높일 수 있는 방법이 요구된다. 이에 본 논문의 4장에서 적응적 트라이맵을 사용하여 품질을 높이는 새로운 방법을 제안한다.

그림 37. 객체 합성 결과

4.2. 얼굴 합성

제안 시스템의 얼굴 합성은 원본 영상에서 사람의 얼굴을 자동으로 추출하며, 사용자에게 입력받은 변경할 얼굴이미지를, 원본영상의 추출된 얼굴에 특징점 매칭을 통해, 3차원 모델을 위치시킨 후, 텍스처 매핑을 해서, 얼굴을 변환시킨다.

다음 그림 38은 얼굴 합성의 전체 처리 순서를 나타낸다.

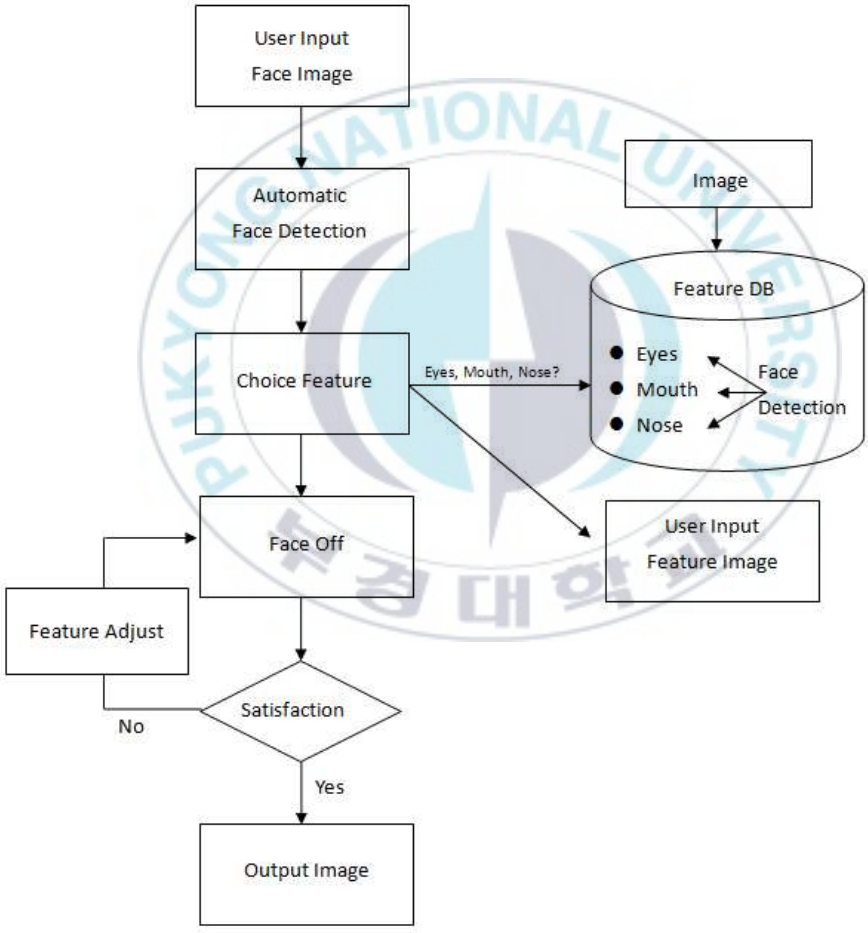


그림 38. 얼굴 합성 처리 순서

제안 시스템에서는 얼굴을 자동으로 검출(Automatic Face Detection)하는 과정에서 Viola와 Jones가 제안한 Haar-like Feature 알고리즘[39]을 적용하여 구현하였다. 이 알고리즘은 특히 얼굴의 특징점인(눈,코,입)을 검출하기 위한 용도로 많이 사용되어진다. 제안 시스템과 같이 실시간(real-time)으로 움직이는 영상에서 매핑시키기 위해서는 빠른 연산속도가 필요한데, 이 기법은 얼굴 검출시에 빠른 속도를 보장하기 때문에 제안시스템의 구현 목적에 부합한다. 또한 이 알고리즘은 적분 영상을 이용해서 연산하며, 환경 요소등의 잡음에 강인하고, 빠른 연산 속도로 얼굴 성분 검출할 수 있는 특징을 가진다.

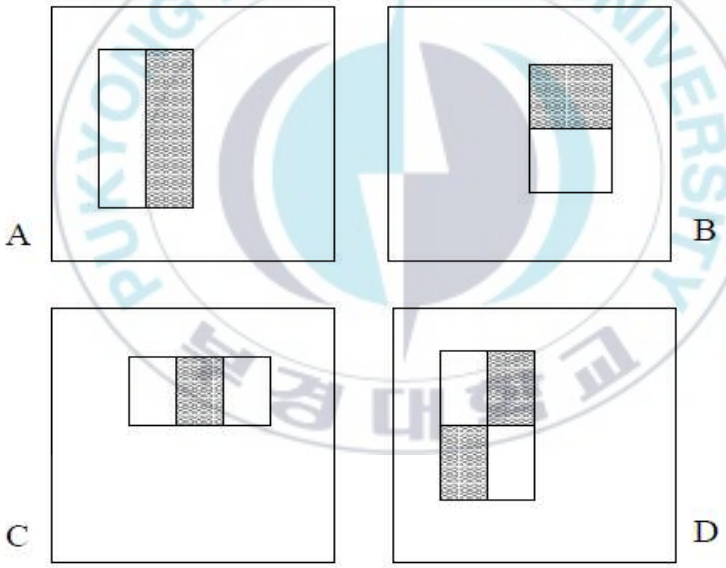


그림 39. haar-like 사각 특징(feature)

위의 그림 39 는 Haar-like[47]의 사각 특징(rectangle feature) 4가지를 나타낸 것으로 A, B는 이미지의 엣지(edge)를, 아래의 C, D는 이미지의 선(line)을 검출하기 위해 쓰이는 특징(feature)들이다. 이 사

각 특징들은 검은색 사각형 영역과 흰색 사각형 영역에서의 각각의 픽셀들 합의 차이를 이용해서 특징을 검출해낸다. 아울러 적분 영상을 사용하기 때문에 연산 속도가 빠르다.

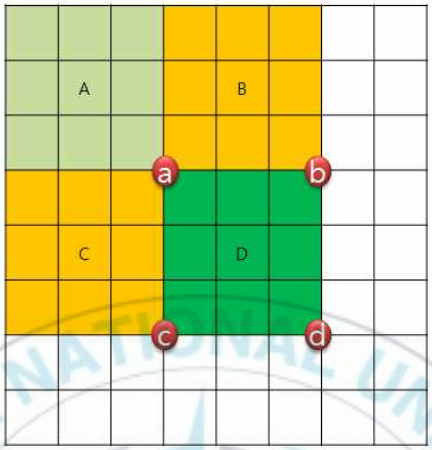


그림 40. 적분 영상의 정의

그림 40은 적분 영상(intergral image)을 원본 영상(original image)과 동일한 크기의 행렬로 표현한 2차원 참조테이블(lookup table)을 나타낸다. 여기서 a는 A영역의 합을, b는 A영역과 B영역의 합을 나타낸다, c는 A영역과 C영역의 합을 나타내며, d는 A,B,C,D영역의 총합, 즉 입력이미지의 총합을 나타낸다. 그러므로 D영역의 총합을 구하려면, (a+d)-(b+c) 계산으로 구해진다.

10	15	7	9	5	6	4	2
9	32	65	45	12	7	5	2
7	24	66	65	41	34	4	7
9	11	70	89	44	37	42	11
32	78	91	78	48	65	15	24
64	12	89	58	65	45	37	9
1	1	96	89	56	48	59	3
2	3	45	65	44	71	57	4

10	25	32	41	46	52	56	58
19	66	138	192	209	222	231	235
26	97	235	354	412	459	472	483
35	117	325	533	635	719	774	796
67	227	526	812	962	1111	1181	1227
131	303	691	1035	1250	1444	1551	1606
132	305	789	1222	1493	1735	1901	1959
134	310	839	1337	1652	1965	2188	2250

그림 41. 원본 영상과 적분 영상의 픽셀합 예시

그림 41 의 왼쪽은 원본 영상을 나타내고, 오른쪽은 적분 이미지를 나타낸다. 원본 영상의 픽셀의 총합은 적분 영상에서 A영역(값:235)과 D영역(값:1444)을 더한 합에서 B영역(값:459)과 C영역(값:691)을 뺀 것과 동일하다. 이와 같이 적분 영상으로 손쉽게 특정한 영역의 픽셀값의 총합을 구할 수가 있다. 이는 Haar-like 사각 특징을 적용하는 적분 이미지를 의미한다. 다음 식(1) 은 적분 영상을 구하는 계산식이다.

$$ii(x,y) = \sum_{y'=0}^{y} \sum_{x'=0}^{x} i(x',y') \quad \text{.....(1)}$$

식 (1)에서 $i(x',y')$원본 영상을 나타내며, $ii(x',y')$은 적분 영상을 나타낸다. 적분 영상은 원본 영상의 시작점 $(0,0)$부터 마지막점 (x,y)까지 모든 픽셀들의 합을 가질 수 있다.

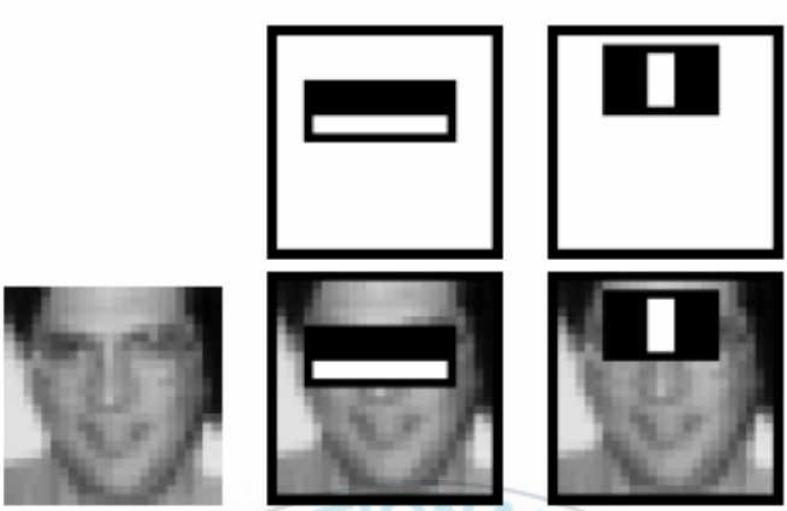


그림 42. Haar-like 사각 특징을 적용한 예시

그림 42에서 보는바와 같이 입력 얼굴 영상에 그림 39의 Haar-like 사각 특징을 적용하여 검은색 영역과 흰색 영역의 픽셀값들의 합을 식 1을 이용해서 계산한 후, 두 합 사이의 차를 계산해서 각 성분의 특징값을 얻어 낼 수 있다. 다시말해, 그 차이가 임계치를 넘으면 이는 사람의 얼굴 Haar-like 사각 특징이 된다는 것이다. 이는 특정 분포에 따른 얼굴의 밝기 차이는 거의 없다는 특성에 기인한 것이다. 이와 같은 얼굴 검출 과정을 거쳐, Face Off 단계에서의 최종 얼굴 합성 처리는 그림 43과 같은 과정을 거쳐 3차원 모델에 텍스처 매핑 작업 후, 객체합성과 마찬가지로 이미지의 경계선을 부드럽게 처리 히스토그램에 기반한 뽀아송 이미지 매팅을 적용하여 합성시킨다 .

그림 43. 3차원 Reconstruction

그림 44. 3차원 매핑

1. 원본 이미지 2. ISO 맵 3. 변경 이미지

그림45. Face Swap 과정

얼굴 합성 시스템에 객체 합성시와 마찬가지로 다음과 같은 Poisson 방정식으로 정의하여 마지막 작업을 한다.

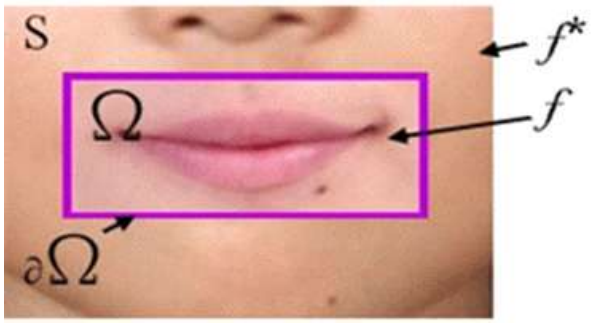
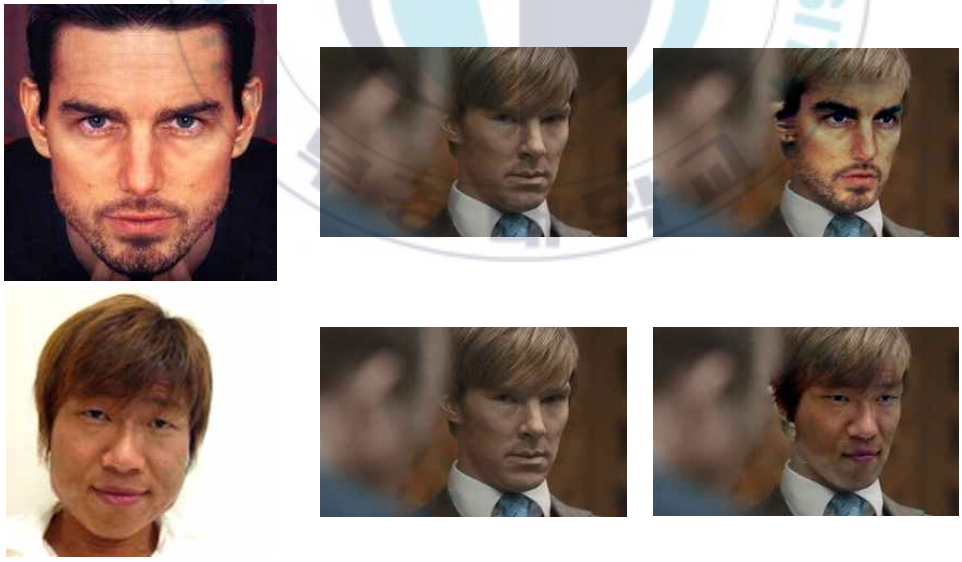


그림 46. 얼굴 합성시 Poisson 매팅

$$\partial\Omega = \{p \in S \setminus \Omega : N_p \cap \neq 0\}, F|\Omega = \{f_p, p \in \Omega\} \quad \text{.....(1)}$$

$$\frac{\min}{f|\Omega} \sum_{\langle p,q \rangle \cap \Omega} (f_p - f_q - v_{pq})^2, with\ f_p = f_p^*, \text{for } all\ p \in \partial\Omega \quad \text{.....(2)}$$



1. 입력 이미지 2. 원본 영상 3. 합성 영상

그림 47. 얼굴 합성 결과

4.3. 배경 합성

배경을 합성하기 위해서는 먼저, 원본 영상에서 변경하고자 하는 부분의 영역을 분류해야한다. 제안 시스템에서는 사용자가 몇 번의 간단한 마우스 조작만으로도 추출할 수 있도록 구성하였으며, Watershed Segmentation 기법을 적용하여 구현하였다. 아래 그림 48은 작업하고자 하는 원본 이미지를 나타내며, 그림 49는 교체될 배경 영역을 분류하는 과정으로서 마우스의 간단한 드래깅 조작을 표시한 것이다. 이와 같은 조작을 통해서 그림50 과 같이 영역이 분류된 이미지를 구해낼 수 있다. 이렇게 구해낸 이미지를 다시 그림 51과 같은 이진화된 마스크로 만들고, 이 마스크를 그림 52와 같은 변경하고자 하는 배경 그림에 씌워서, 최종적으로 그림 53과 같은 원본 영상과 배경이 합성되어진 영상을 만들어 낸다.

그림 48. 원본 영상

그림 49. 마우스 드래그 지정

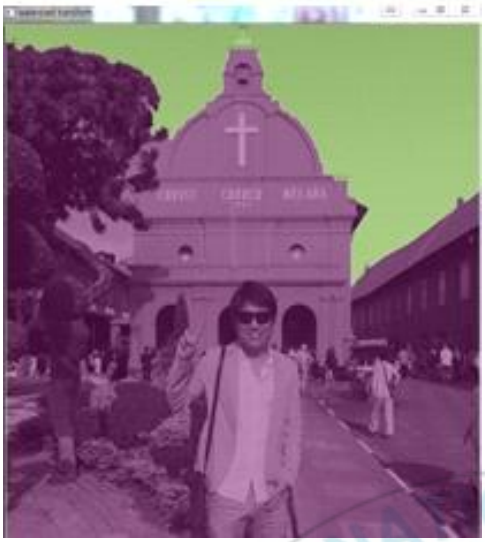


그림 50. 하늘 영역 분류

그림 51. 마스크 적용

원본 영상에서 마우스로 구분하고자 하는 영역을 적당히 드래그 하면, 영역이 분류된다.

그림 52. 배경 영상

그림 53. 배경 합성 결과

위 그림53은 배경의 구름이 있는 하늘로 합성한 결과물을 나타낸다.

제안 시스템에서는 배경을 합성하기 위해서, 영역을 구분해내는 알고리즘으로 Lantuejoul과 Beucher이 제안한 Watershed 알고리즘[48]을 적용하여 구현하였다. 이 기법은 영역 확장을 이용하여 경계선을 구해내는 방식으로 주로 영상 분할을 하기 위해서 많이 사용되어진다.

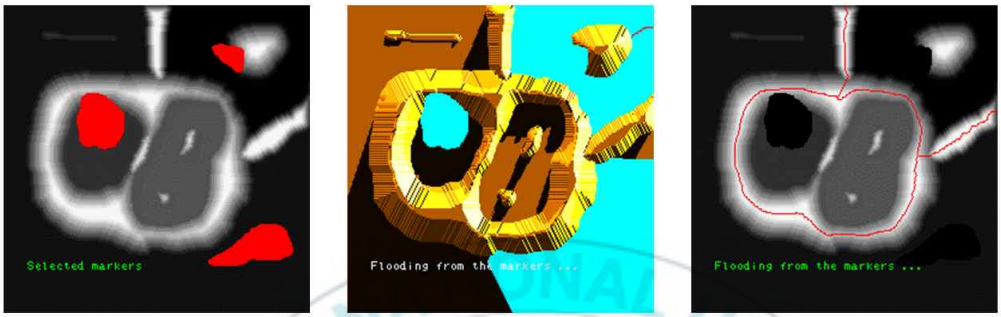


그림 54. Watershed Segmentaion 과정

Watershed 알고리즘은 그림 53과 같이 이미지의 픽셀값을 지형과 같은 면(topographic surface)로 간주하고, 지형상에서 가장 낮은 지역에서부터 물을 채워 영역을 확장하는 방식으로 전개한다. 제안 시스템에서는 초기 시작점이 되는 시드(seed)를 마우스의 드래깅으로 지정하도록 구현하였다. 영역은 밝기값의 기울기의 고도에 따라서 점점 영역을 확장해가면서, 이 영역이 마치 저수지에 물이 차듯 점점 차오르면서 봉우리 지점에서 멈추면, 다른 인접한 영역의 물이 차오르면서 봉우리 지점에서 만나게 된다. 이 때 두 영역, 물이 차오른 담수 지역은 경계선을 만들어진다. 이러한 방법으로 시드에서 최소값으로 시작한 영역들은 결국 독립된 영역으로 구분된다. 이와 같이 watershed 알고리즘은 시드를 중심으로 확장해 나가기 때문에 국부적 오류가 범람하지 않고 차단되어서, 전역 오류로 확산되는 걸 막아 주는 장점을 가지고 있다.

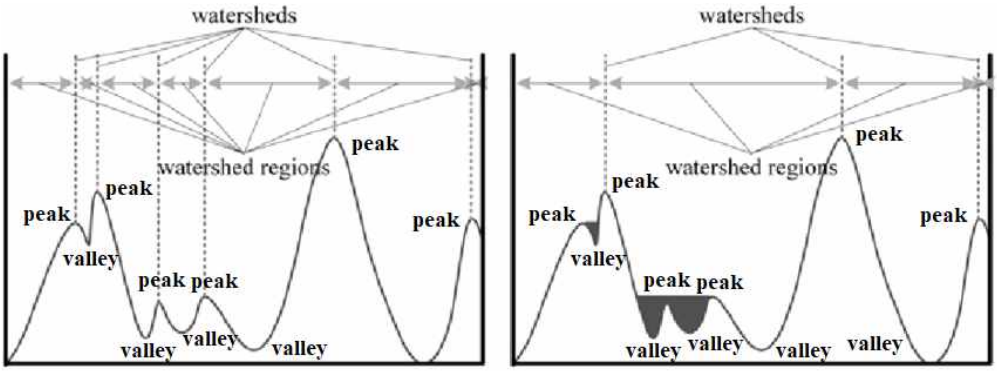


그림 55. Watershed 알고리즘

그림 55는 측면에서 바라보는 지형면을 표현한다. 계곡(valley)은 지형의 가장 낮은 곳을 의미하고, 꼭대기(peak)는 지형의 가장 높은 곳을 의미한다. 계곡(valley)에서부터 서서히 물을 채워 가면서, 꼭대기(peak)까지 이르면 범람(flooding)하는 위치를 찾는 방식으로 진행된다. 이때 초기 시작 지점을 나타내는 시드(seed)와 가장 가까운 물을 채우는 가장 낮은 지형의 계곡(valley)을 같은 영역으로 간주한다. 이때 시드(seed)는 지정해주는 작업이 필요하다. 이렇게 시드와 가장 낮은 지형이 결정되면, 다음 과정으로 점점 영역을 확장해 나가게 된다. 이렇게 영역 확장 과정에서는 우선순위가 필요한데, 이 우선 순위는 식(1)과 같이 시드(seed)의 인접 픽셀의 밝기(inetensity)값의 기울기(gradient)로 정의된다.

$$p(x,\ y) = \nabla f(x,y) \quad \text{.....(1)}$$

$$f(q) = \min\ r \in N_G(p)\{f\ (r) \mid f(r) < f(p)\} \quad \text{.....(2)}$$

$$f(q) = f(p) \text{ and } d(p) = d(q) + 1 \quad \text{.....(3)}$$

식(2)에서 $f(p)$ 는 픽셀의 현재 위치 p 에서의 기울기를 나타내고,

$N_G(p)$ 는 인접 픽셀들을 의미한다. 식(2)와 식(3)의 조건을 맞으면 같은 영역으로 간주한다. 여기서 기울기가 작을수록 고도가 낮기 때문에 영역을 확장시 우선순위를 갖게 된다. 이런 우선순위를 기준으로하여 초기시작점 시드(seed)와 계곡(valley)에서부터 낮은 고도는 빠르게, 높은 고도는 천천히 영역을 확장하면서, 꼭대기(peak)에 다다르게해서 경계선을 구해내게 된다. 이렇게 구한 경계선을 따라서 최종적으로 영역을 분할되어 구분되어진다.

구현 시스템의 이와 같은 알고리즘과 마우스 드래그로 시드(Seed)를 설정하는 사용자 인터페이스에서는 결합하여, 영역을 확장하면서 배경과 이외의 영역을 구하도록 구현 하였다.

IV 적응적 트라이맵을 사용한 영상 합성

1. 트라이맵 기반 영상 합성

대부분의 이미지 매팅 기법과 마찬가지로 Poisson 매팅 기법 또한 트라이맵을 사용하는 영상 합성 기법이다. 그러므로, 합성 연산을 수행하기 위해서는 그림 56와 같이 입력 전경 이미지와 함께 아래 트라이맵이 주어져만 한다. 하지만 복잡한 이미지에서 세밀하게 트라이맵을 설정하는 문제는 쉬운 작업이 아니다.

그림 56. 입력 전경 이미지와 트라이맵

구현 시스템에서는 마우스 클릭에 의한 사용자 입력을 통해 알 수 없는 영역(Unknown) 바깥쪽 외부 경계를 설정하는 작업을 통해 사용자가 이미지를 확인하면서 상호작용적으로 트라이맵을 설정하도록 하였다.

그림 57-1과 같이 사용자가 객체에 대강의 경계로 트라이맵 외부 경계를 설정해주는 것만으로도, 그림 57-3과 같이 자연스러운 결과를 얻을 수 있다.

1.사용자 트라이맵 설정 2.원본 영상 3.합성 영상

그림 57. 사용자 트라이맵을 통한 Poisson 합성 결과

그림 58 같은 방법의 실험 결과이다. 마찬가지로 결과가 자연스러운 것을 확인 할 수 있다.

1. 객체 추출 2.원본 영상 3.합성 영상

그림 58. 사용자 트라이맵을 통한 Poisson 합성 결과 2

하지만 그림 57과 그림 58의 두 가지 경우 모두 바닷물, 초원과 같이 비슷한 전경과 배경으로서 색상값이 차이가 크지 않고, 또한 전경과 배경은 각각 밝기값의 기울기(gradient)가 평활(smooth)하다는 공통적인 특징이 있다. 그래서 본 연구에서는 두 가지 경우처럼 최적이 아닌 일반적인 상황에서도 합성 품질을 보장하기 위하여, 다양한 이미지의 조건에 적응적으로 트라이맵을 사용하는 방법을 제안한다.

2. 적응적 트라이맵 기반 합성 및 실험 결과

아래 표3과 같이 합성 대상 이미지가 상황 0과 같이 최선의 조건을 가지는 경우 1가지와, 최선이 아닌 조건을 가지는 경우를 4가지로 정리하고, 적응적 트라이맵 사용을 위한 접근법을 제시한다. 아래 표3의 푸른색 배경은 좋은 조건을 의미하고, 반대로 붉은색은 나쁜 조건을 의미한다.

경우 (Case)	합성 대상 이미지의 조건			접근법(Approach) - 적응적 트라이맵 조정
	전경	배경	색상 차	
0	단순	단순	작다	방법1-사용자 제어(라소 도구)
1	단순	복잡	작다	방법1-사용자 제어(라소 도구), 방법2-외부 경계 조정
2	복잡	복잡	작다	방법3-외부 경계 재설정
3	단순	복잡	크다	방법3-외부 경계 재설정, 방법4-내부 경계 추가 설정
4	복잡	복잡	크다	방법3-외부 경계 재설정, 방법4-내부 경계 추가 설정, 방법5-트라이맵 국부 지역 재설정

[표 3] 합성 대상 이미지의 조건과 접근법

2.1. Case 1 - 전경(단순), 배경(복잡), 색상차이(작음)

그림 59과 같이 Case 1의 경우는 배경이 복잡하다는 나쁜 조건을 가지고 있다. 즉 영상의 밝기값의 기울기가 평활하지 않다는 의미다. 그림 59의 4의 합성 결과는 그림59의 3가 같이 사용자 입력에 의한 트라이맵을 설정했으때의 결과이다.

1. 배경

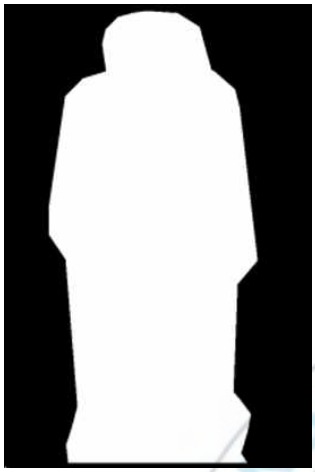
2.전경

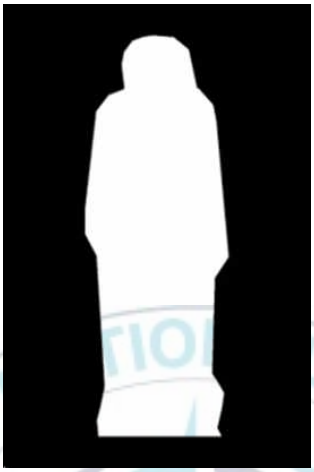
3.트라이맵

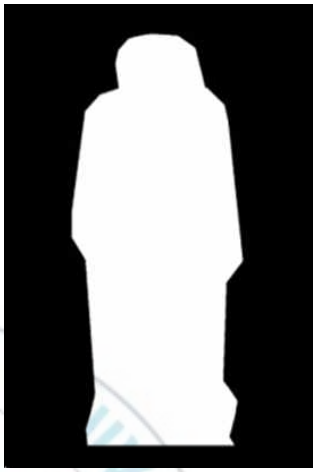
4.합성 결과

그림 59. Case 1 합성 과정

위의 결과를 개선하기 위해서, 트라이맵의 외부 경계를 조정한다. 이때 사용자의 추가적인 사용자의 개입은 필요치 않는다. 자동으로 생성되어 제공되는 합성 결과를 선택하기만 하면 된다.


1. 트라이맵 조정 1


2. 트라이맵 조정 2


3. 트라이맵 조정 3


4. 합성 결과 1


5. 합성 결과 2


6. 합성 결과 3

그림 60. 트라이 맵 조정 후 합성 결과

그림 60의 경우 5의 합성 결과2가 가장 최선의 결과로 나타난다.

그림 61은 전체 화면으로 나타낸 Case 1의 합성 결과이다.

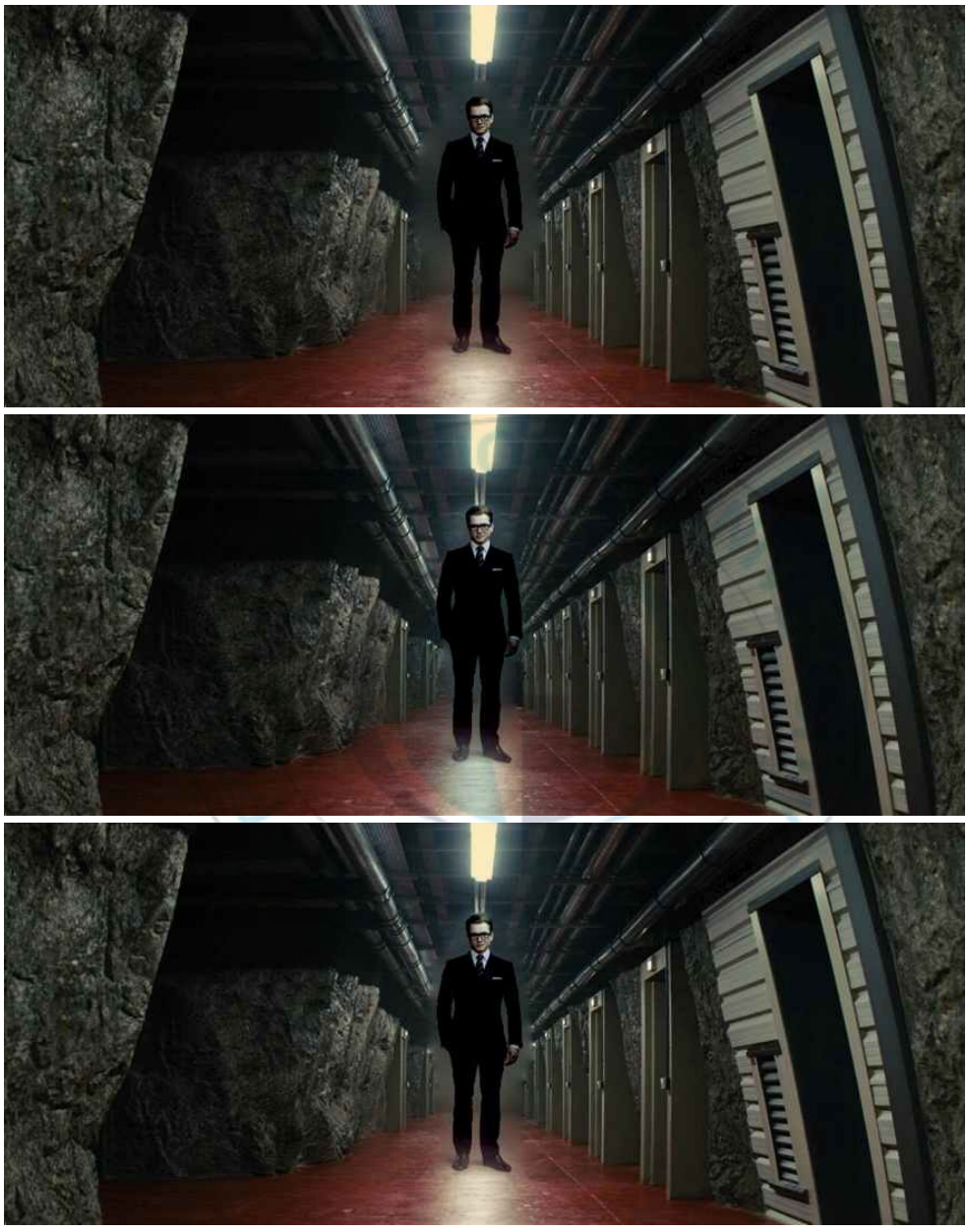


그림 61. Case 1 합성 결과 전체 화면

2.2. Case 2 - 전경(복잡), 배경(복잡), 색상차이(작음)

그림 62와 같이 Case 2의 경우는 Case 1과 마찬가지로 배경이 복잡하다는 나쁜 조건을 가지고, 동시에 전경도 복잡하다는 나쁜 조건을 가진다. 그림62-4의 합성 결과는 방법2-외부 경계 조정을 적용하다. 하지만 머리와 다리 좌측 부분의 합성이 매끄럽지 못하다는 것을 확인할 수 있다. 그래서 이와같은 Case 2의 경우는 그림 63과 같이 트라이맵의 외부 경계를 재설정하는 새로운 방법을 제시한다.

1. 배경

2.전경

3.트라이맵

4.합성 결과

그림 62. Case 2 합성 과정과 합성 결과(방법2 적용)

외부경계 재설정(방법3)은 그림 63의 1과 같이, 배경 합성에서 쓰였던 Watershed 영역 분할 기법을 적용하여, 사용자 컨트롤로 트라이맵의 외부 경계를 재설정하는 작업이다. 이는 사용자제어(방법1:라쏘도구)로 대강의 경계를 따는 작업보다는 조금 더 많은 조작이 필요하다.


1.사용자 컨트롤


2. 영역 분할 과정


3. 외부 경계 재설정


4. 합성 결과1 (방법2)


5. 합성 결과 2(overlap)


6. 합성 결과3 (방법3)

그림 63. Case 2 합성 과정

그림 63-5의 경우는 트라이맵 경계 재설정 작업 후, 그림 63-6처럼 포아송(Poisson) 매팅으로 처리하지 않고, 단지 오버랩(overlap)한 결과이다. 머리카락 부분과 다리 아래쪽이 어색하다는 것을 확인 할 수 있다.

1. 합성 결과2 (overlap)　　2.합성 결과3 (방법3)

그림 64. Case2 합성 결과 확대 화면

외부 경계 재설정(방법3)을 적용한 합성 결과3의 경우, 외부 경계 조정(방법2) 방식과 마찬가지로 자동으로 매팅된 몇 가지 결과를 그림 65와 같이 자동 추천해주면, 사용자가 만족하는 품질의 결과를 선택할 수 있다. 위 그림 64-2 합성 결과3도 아래 그림65의 4개의 중에서 하나인 3번째 결과이다.

그림 65. Case 2의 방법3 적용 후 합성 결과

그림 66. Case 2의 방법3 합성 결과 전체 화면

위의 그림 65와 그림 66은 그림 67에서 확대한 결과를 보고 차이점을 확인 할 수 있다.

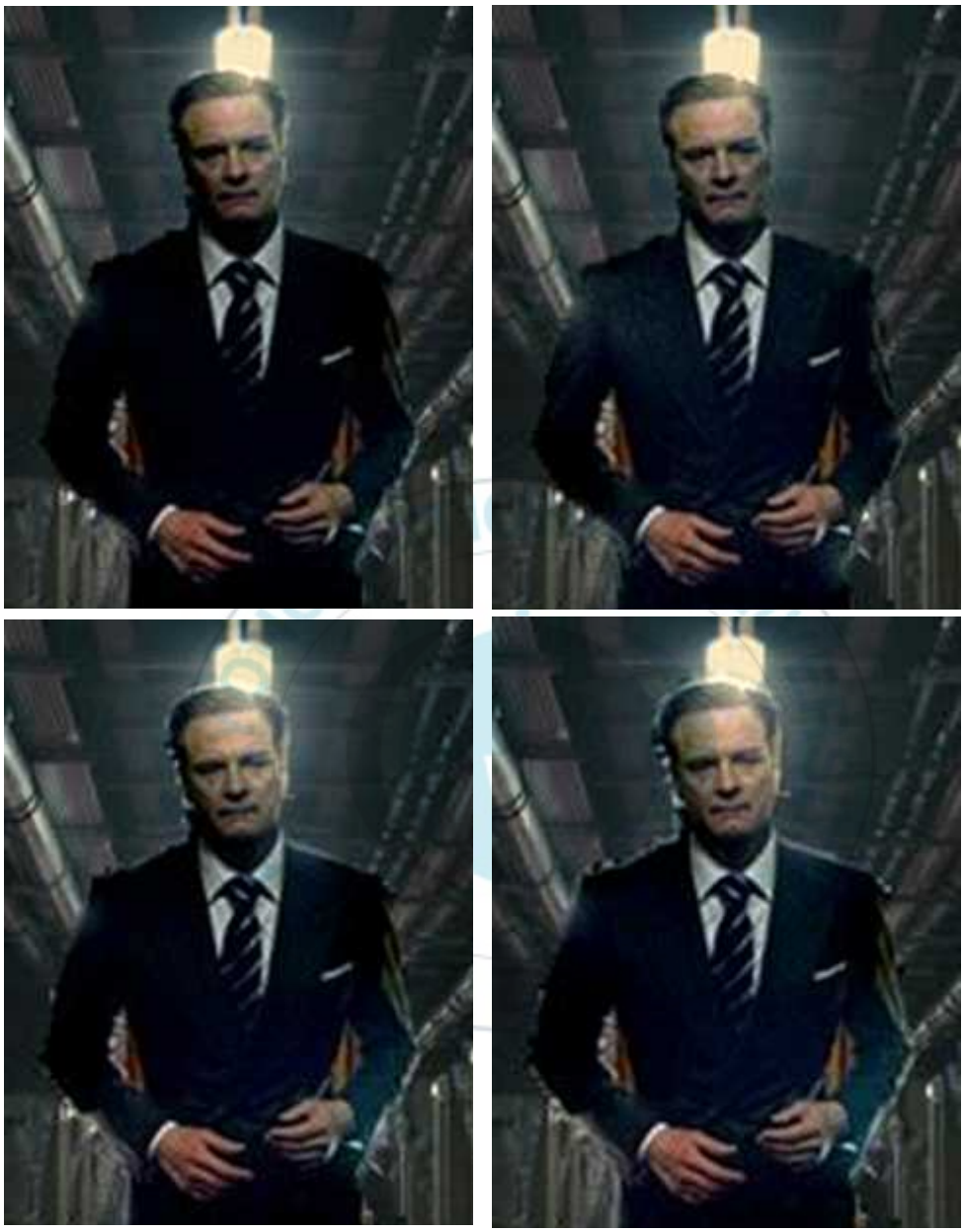


그림 67. Case 2 의 방법3 합성 결과 확대 화면

2.3. Case 3 - 전경(단순), 배경(복잡), 색상차이(큼)

그림 68과 같이 Case 3의 경우는 배경이 복잡하다는 나쁜 조건을 가지고, 아울러 전경과 배경의 색상값의 밝기 차이가 크다는 새로운 나쁜 조건을 가진다. 물론 전경이 단순하다는 좋은 조건도 있다.

1.배경 2.전경

그림 68. Case 3 의 배경과 전경

Case 3의 경우 앞에서 제시한 적응적 트라이맵 기법들을 사용하여 합성 결과를 아래 그림 69에서 확인한다.

1. 합성 결과1
트라이맵 조정(방법2)

2. 합성 결과2
외부경계 재설정(오버랩)

3. 합성 결과3
외부경계 재설정(방법3)

4. 합성 결과1 확대

5. 합성 결과2 확대

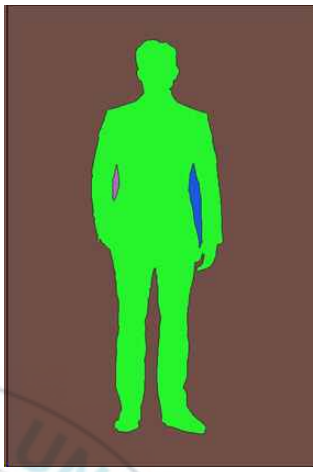
6. 합성 결과3 확대

그림 69. Case 3 합성 결과

위의 그림 69-6의 합성 결과3 에서 신발 부분이 투명해진 것을 볼 수 있다. 이러한 문제를 해결하고자 아래 그림 70과 같이, 트라이맵의 내부 경계를 추가 설정하는 새로운 방법을 제시한다(방법4). 이 기법은 아래 그림 70-1과 같은 추가적인 사용자의 조작 필요로 하지 않는다.

왜냐하면 이미 외부경계 재설정(방법3) 과정에서 사용자가 조작을 해놓았기 때문이다.

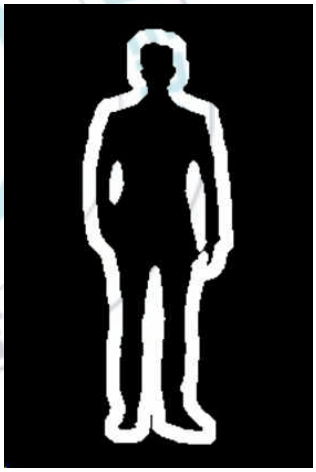




1. 트라이 맵 외부경계 재설정을 위한 사용자 컨트롤

2. 내부 경계 추가 설정.

3. 외부 경계 조정1



4. 외부 경계 조정2

그림 70. 내부 경계 추가 설정

그림 71-3 은 내부 경계 추가 설정 후 합성 결과를 나타낸다.

1. 합성 결과1
외부경계 재설정(오버랩)

2. 합성 결과2
외부경계 재설정(방법3)

3. 합성 결과3
내부경계 추가설정(방법4)

그림 71. Case 3 의 합성 결과

2.4. Case 4 - 전경(복잡), 배경(복잡), 색상차이(큼)

그림 72와 같이 Case 4의 경우는 배경과 전경이 복잡하고, 아울러 전경과 배경의 색상값의 밝기 차이가 크다는 모든 나쁜 조건을 다 가진다.

1.배경 2.전경

그림 72. Case 4 의 배경과 전경

그림 73은 Case 4의 합성 결과를 나타낸다.

1. 합성 결과1
외부경계 재설정(오버랩)

2. 합성 결과2
외부경계 재설정(방법3)

3. 합성 결과3
내부경계 추가설정(방법4)

그림 73. Case 4 의 합성 결과

Case 4의 경우는 Case 3의 경우와 동일하게 제시한 다양한 적응적

트라이맵 사용 기법을 통하여 합성을 진행한다. 다만. 그림 74와 같이 합성 후 지역 포아송(Poisson) 매팅을 재적용하여, 오류를 보정한다.

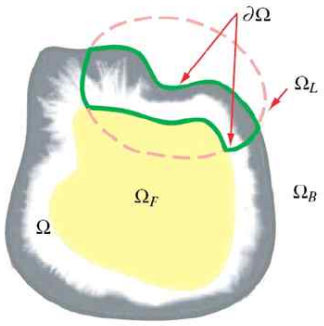
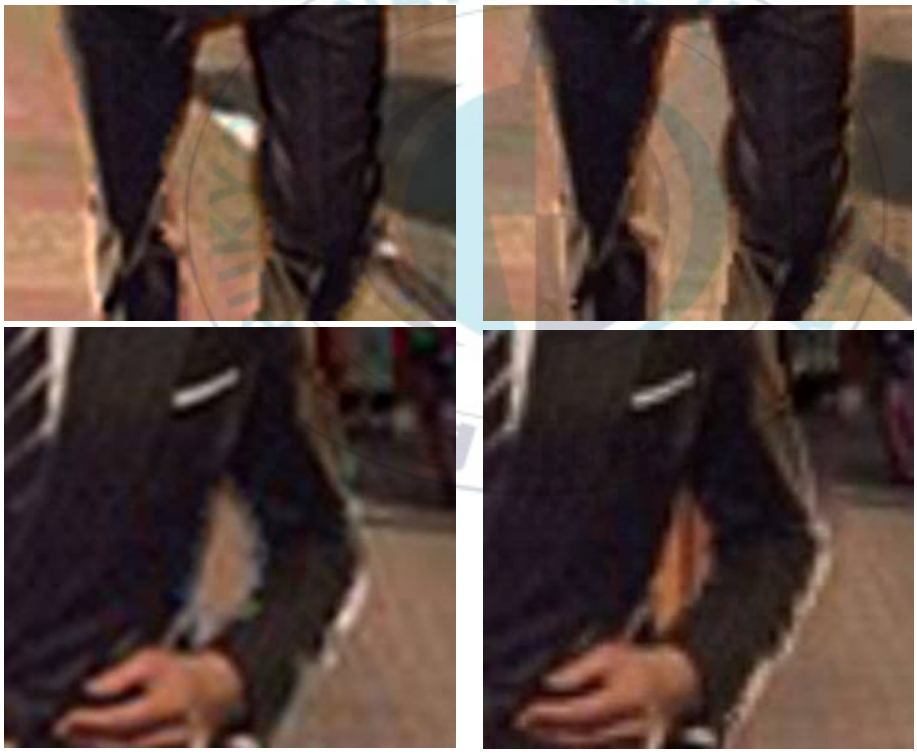


그림 74. 지역 Poisson 매팅



1. 합성 결과1
내부경계 추가설정 (방법4)

2. 합성 결과2
트라이맵 국부 지역 재설정 (방법5)

그림 75. Case 4 의 합성 결과(방법5)

V 상용 시스템과의 비교

연구에서 제안하는 시스템을 구현결과를 비교하기 위해서, 먼저 가장 우수하다고 알려진 Adobe사의 After Effect를 사용하여 원본 영상에 대해 객체합성, 얼굴합성, 배경합성을 각각 수행하여, 실험 영상을 재생성하였다. 이는 제안 시스템 구현 성능 분석과 비교를 위한 작업이다. 콘텐츠 제작을 위해서는 숙련된 기술을 위한 조작툴에 대한 학습이 필요하며, 사용자 개인의 숙련도에 따라서 결과 산출물의 차이가 많이 난다. 본 연구에서는 우수한 성능을 나타내는 저작도구와 같은 기능과 비교하여, 사용자 조작편의 용이성에 중점을 두고 시스템을 비교 분석하고자 하였다. 이에 상용 저작 도구 조작 방식과 결과물을 정리하고, 아울러 제안 시스템의 결과물도 정리한다.

1. 저작도구를 이용한 합성 실험

1.1. 객체 합성

After Effect를 이용한 실험에서, 움직이지 않는 객체(예:테이블 위 음식)에 대해서 합성 실험을 하였다. 그림 76의 왼쪽은 원본 영상이고, 오른쪽은 합성 영상을 나타낸다. 위에서 아래쪽으로 갈수록 시점(카메라의 위치)가 변하고 있음을 나타낸다. 그림 76에서 보듯이 장면 전환이 되지 않은 유사 이미지로 구성된 영상으로, 참조구간에서 제한적인 작업이 이루어졌다.

그림 76. 객체 합성 결과 (저작 도구)

위 그림76 에서 After Effect를 이용한 작업 절차는 다음과 같다.

① 합성을 시도할 영상과 합성할 사물사진을 불러온다. 이때 픽셀사이즈는 유사하게 맞춘다.
② 동영상 위에 사진을 올린 후 Motion Tracking 효과를 영상에 적용하여 움직임을 딴다.
③ 움직임을 딴 것을 사진에 Apply해준다.
④ Extract 효과를 사용하여 영상의 배경을 지운다.
⑤ 적절히 배경값을 조절 후 사진의 Brightness, Contrast를 조정한다.
⑥ 자연스러움을 위해 객체에 drop shadow 효과를 적용하여 그림자를 준다.
⑦ 합성의 어색함을 부드럽게 하기 위해 Matte Choker를 합성경계 부분에 적용한다.
⑧ 원본영상과 비슷한 색감을 위해 Curves 를 이용하여 RGB값을 조정한다.

이와같이 복잡한 작업프로세스로 숙련된 전문조작기술이 필요하다.

1.2. 얼굴 합성

얼굴 합성 실험의 경우, Photoshop, Mocha, After Effect 3가지 툴을 사용하여 아래 그림 77 과 같이 영상을 재생성 하였으며, 작업단계는 아래와 같다.

① photoshop을 이용, 합성할 사진의 얼굴을 자른다.
② mocha에서 tool을 사용해 동영상 얼굴 움직임을 딴다.
③ export date를 이용하여 움직임의 소스를 복사한다.
④ after effect에서 사진 위에 복사한 소스를 붙여 넣는다.
⑤ 사진과 영상 속 인물의 얼굴 크기를 조절한다.

이와 같이, 다양한 툴을 활용 할 수 있어야 한다.

그림 77. 얼굴 합성 결과 (저작 도구)

1.3. 배경 합성

배경 합성의 경우 그린 스크린을 적용 기법과 저작도구의 모션트랙킹 기법을 사용하여 실험하였다.

그림 78. 배경 합성 결과 1 (저작 도구)

첫 번째, 그림 78의 그린 스크린 기법의 작업단계는 아래와 같다.

① 합성을 시도할 영상과 합성할 사진을 불러온다. 이때 픽셀사이즈는 유사하게 맞춘다.

② 동영상 위에 사진을 올린 후 Motion Tracking 효과를 영상에 적용하여 움직임을 딴다.

③ 움직임을 딴 것을 사진에 Apply해준다.

④ 영상을 Duplicate한 후 Color Keying 효과를 사용하여 영상의 이미지를 지운다.

⑤ 적절히 Keying 오차 값을 조절 후 사진의 Brightness, Contrast를 조정한다.

⑥ 합성의 어색함을 부드럽게 하기 위해 Matte Choker를 합성경계 부분에 적용한다.

⑦ 원본영상과 비슷한 색감을 위해 Curves 를 이용하여 RGB값을 조정한다.

그림79 는 두 번째 기법의 Motion Tracking 기능을 이용하여 하늘의 배경을 합성한 결과를 보여준다.

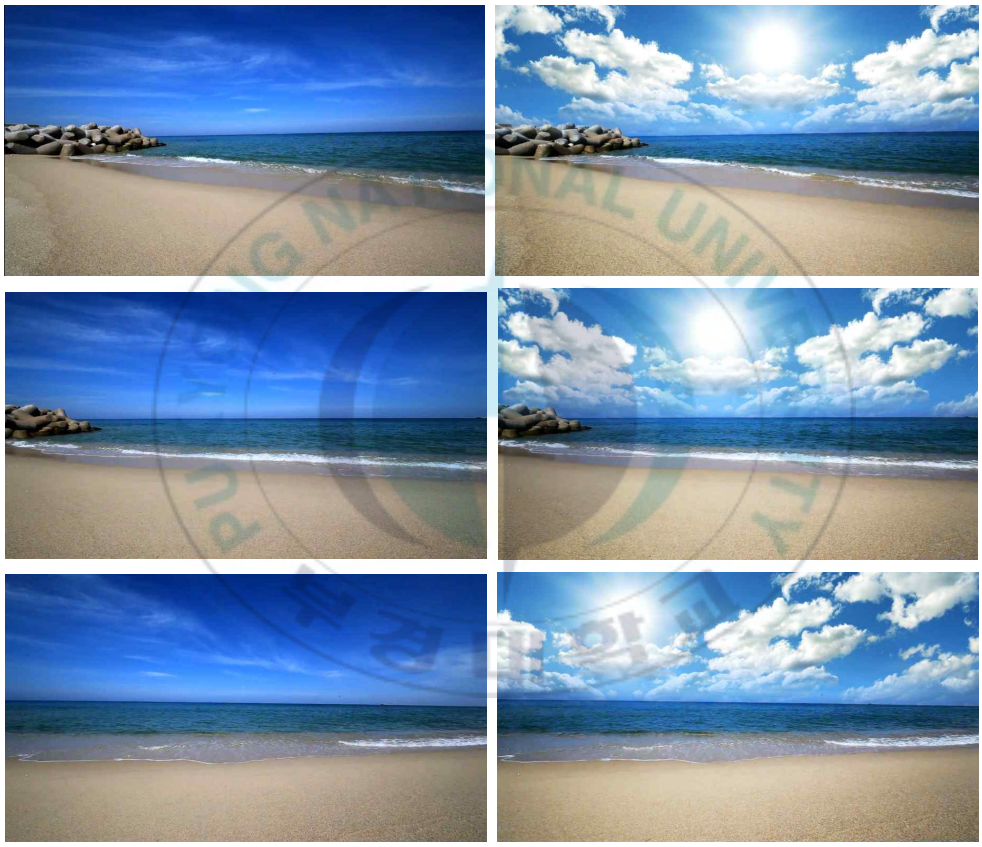


그림 79. 배경 합성 결과 2 (저작 도구)

그림79의 배경합성 실험2 에 적용한 작업 단계는 아래와 같다.

①동영상과 합성할 배경 사진을 준비한다.
②effect-transition-linear wipe feather(페더)로 합성할 배경 경계부분 색 번지게 한다.
③동영상의 일부분을 motion tracking을 적용한다.
(motion tracking시 option에서 enhance 체크로 변경)
④움직임을 사진에 적용한 상태로 위치를 맞춘다.
⑤사본을 만들고 Keying 작업을 통해 합성하고자 하는 부분에 색을 없앤다.
⑥새로운 레이어를 추가해 명도와 채도를 조정 후 렌더링한다.

2. 제안 시스템을 이용한 합성 실험

2.1. 객체 합성

제안 시스템에서의 얼굴 합성 결과는 아래 그림 80 과 같다.

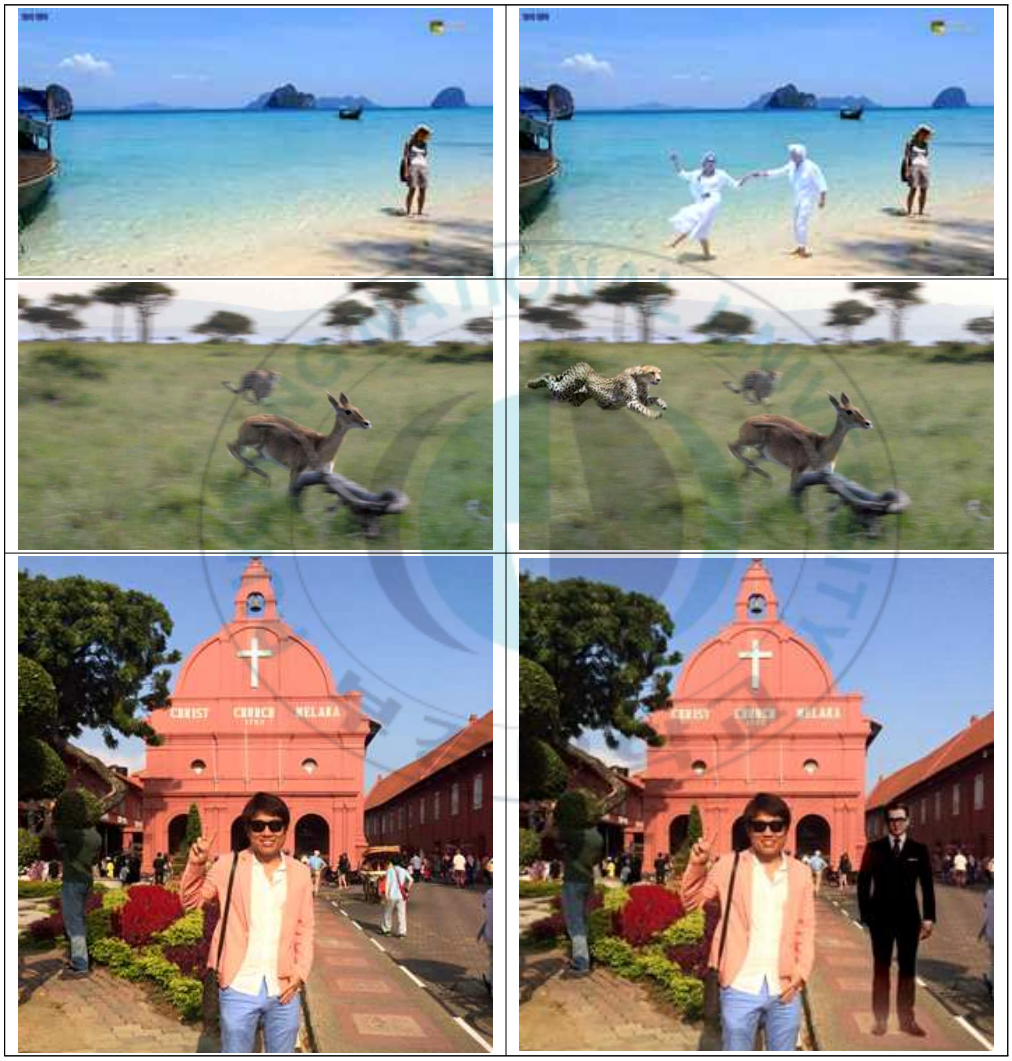


그림 80. 객체 합성 결과

2.2. 얼굴 합성

제안 시스템에서의 얼굴 합성 결과는 아래 그림 81, 82 와 같다

그림 81. 얼굴 합성 결과 1 (한국인)

그림 82. 얼굴 합성 결과 2 (서양인)

그림 83의 두 얼굴을 합성한 결과는 그림84과 85와 같다.

그림 83. 얼굴 합성 소스

그림 84. 얼굴 합성 결과 3 (영화)

그림 85. 얼굴 합성 결과 4 (영화)

2.3. 배경 합성

구현 시스템에서의 배경 합성 결과는 아래 그림 86 와 같다.

그림 86. 배경 합성 결과 1 (광고)

그림 87. 배경 합성 결과 2 (광고)

VI 결론

동영상 컨텐츠 플랫폼의 급격한 성장에 따라 콘텐츠 제작에 있어서도 사용자의 의견이 반영되는 맞춤형 영상 수요가 증가하였다. 그러나 아직까지도 영상을 제작할 때 비숙련자가 상용 멀티미디어 영상 편집 저작 도구들을 활용하여 작업하려면 상당히 불편하다. 왜냐하면 영상 합성에서 자연스러운 결과를 얻으려면 많은 수작업이 필요하기 때문이다. 그래서 본 논문에서는 사용자 개개인이 영상 제작 요구를 충족시키기 위해 우리는 조작을 최소화하면서 상용 시스템과 유사한 효과를 낼 수 있는 합성 영상 재생성 시스템을 제안하였다. 객체 합성은 경계는 반자동으로 Poisson 이미지 매팅 기술로 부드럽게 처리하고자 하였고, 얼굴 합성의 경우 Haar-like 특징 알고리즘을 통해 원본 비디오의 얼굴 움직임이 자동으로 추적한 후, 새로운 얼굴 이미지로 교체하였다. 배경 합성의 경우 Watershed Segmentation을 통한 간단한 조작으로 전경과 배경을 구분하여 변경하였다. 특히 객체합성의 경우 합성 대상 이미지의 복잡성에 따라 적응적으로 트라이맵을 적용하는 새로운 기법을 제시하여 합성 품질을 높이고자 하였다. 본 논문의 제안 시스템이 개인 콘텐츠 제작자에게 도움이 되기를 기대해본다.

참고 문헌

[1] 유초롱, 박소영, 조기성 “소셜 TV 서비스에 대한 연구.” 한국통신학회 학술대회논문집,,267-268(2013)

[2] 황지은, 이상희, 김한결, 서지원, 민준기 “사용자 참여형 모바일 증강현실 콘텐츠 제작 체계에 대한 기초 연구.” 한국HCI학회 학술대회, 921-923(2012)

[3] 이성희, 김동칠, 정광수 “스마트 TV 중심의 콘텐츠 공유 서비스를 위한 SVC 기반의 전송률 조절 기법.” Telecommunications Review. vol.22,no.2,pp. 258-270(2012)

[4] 김민수, 정광수 “시청자 반응 및 의도 기반의 미디어 인터랙션 서비스를 위한 컨텍스트 처리 시스템.” 한국정보과학회 학술발표논문집,1194-1196(2018)

[5] 홍석진, 강은범, 김재면, 안기옥, 이승형, 채옥삼 “LBP와 컬러 히스토그램을 이용한 자동화된 장면 검출 기반 장면단위 동영상 편집 시스템.” 한국정보과학회 학술발표논문집,1725-1727(2014)

[6] Cotsaces, Costas, Nikos Nikolaidis, and Ioannis Pitas. "Video shot detection and condensed representation. a review." IEEE signal processing magazine 23.2, 28-37. (2006)

[7] W. Lin et al. (Eds.), “Multimedia Analysis,” Processing and Communications, Springer(2011)

[8] https://trecvid.nist.gov/

[9] Greg Welch, Gary Bishop, University of North Carolina at Chapel Hill, “An Introduction to the Kalman Filter.” SIGGRAPH (2001), Course 8.(2001)

[10] Deutscher, J., Blake, A., & Reid, I. (2000, June). “Articulated body motion capture by annealed particle filtering.” In Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662) (Vol. 2, pp. 126-133). IEEE.(2000)

[11] Sun, Deqing, Stefan Roth, and Michael J. Black. “Secrets of optical flow estimation and their principles.” 2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, (2010)

[12] Comaniciu, Dorin, Visvanathan Ramesh, and Peter Meer. "Kernel-based object tracking." IEEE Transactions on Pattern Analysis & Machine Intelligence 5, pp.564-575. (2003):

[13] Comaniciu, Dorin, and Peter Meer. "Mean shift analysis and applications." Proceedings of the Seventh IEEE International Conference on Computer Vision. Vol. 2. IEEE, (1999)

[14] Comaniciu, Dorin, and Peter Meer. "Mean shift: A robust approach toward feature space analysis." IEEE Transactions on Pattern Analysis & Machine Intelligence 5 pp.603-619.(2002)

[15] Allen, John G., Richard YD Xu, and Jesse S. Jin. "Object tracking using camshift algorithm and multiple quantized feature spaces." Proceedings of the Pan-Sydney area

workshop on Visual information processing. Australian Computer Society, Inc.(2004)

[16] Hidayatullah, Priyanto, and Hubert Konik. "CAMSHIFT improvement on multi-hue and multi-object tracking." Proceedings of the 2011 International Conference on Electrical Engineering and Informatics. IEEE. (2011)

[17] Hotho, Andreas, et al. "Information retrieval in folksonomies: Search and ranking." European semantic web conference. Springer, Berlin, Heidelberg, (2006)

[18] Smeulders, Arnold WM, et al. "Content-based image retrieval at the end of the early years." IEEE Transactions on Pattern Analysis & Machine Intelligence 12: pp.1349-1380. (2000)

[19] Liu, Guang-Hai, and Jing-Yu Yang. "Content-based image retrieval using color difference histogram." Pattern recognition 46.1: pp.188-198. (2013)

[20] Yu, Hui, et al. "Color texture moments for content-based image retrieval." Proceedings. International Conference on Image Processing. Vol. 3. IEEE, (2002)

[21] Zhang, Dengsheng, and Guojun Lu. "Review of shape representation and description techniques." Pattern recognition 37.1: pp.1-19. (2004)

[22] Avidan, Shai, and Ariel Shamir. "Seam carving for content-aware image resizing." ACM Transactions on graphics (TOG). Vol. 26. No. 3. ACM, (2007)

[23] Mishima, Yasushi. "Soft edge chroma-key generation based upon hexoctahedral color space." U.S. Patent No. 5,355,174. 11 Oct. (1994)

[24] C. CORPORATION, “Knockout user guide,” (2002)

[25] Ruzon, M. A., & Tomasi, C. (2000, June). Alpha estimation in natural images. In Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662) (Vol. 1, pp. 18-25). IEEE.(20000)

[26] CHUANG, Yung-Yu, et al. A bayesian approach to digital matting. In: CVPR (2). pp. 264-271.(2001)

[27] P□rez, Patrick, Michel Gangnet, and Andrew Blake. "Poisson image editing." ACM Transactions on graphics (TOG) 22.3 : 313-318.(2003)

[28] Levin, Anat, Dani Lischinski, and Yair Weiss. “A closed-form solution to natural image matting.” IEEE transactions on pattern analysis and machine intelligence 30.2 : 228-242.(2007)

[29] WANG, Jue; COHEN, Michael F. “Optimized color sampling for robust matting.” In: 2007 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, p. 1-8.(2007)

[30] GRADY, Leo, et al. “Random walks for interactive alpha-matting.” In: Proceedings of VIIP. p. 423-429.(2005)

[31] Kwatra, V., Schodl, A., Essa, I., Turk, G., & Bobick, A. (2003, July). “Graphcut textures: image and video synthesis using graph cuts.” In ACM Transactions on Graphics (ToG) (Vol. 22, No. 3, pp. 277-286). ACM.(2003)

[32] Sun, J., Jia, J., Tang, C. K., & Shum, H. Y. (2004, August). "Poisson matting." In ACM Transactions on Graphics (ToG) (Vol. 23, No. 3, pp. 315-321). ACM.(2004)

[33] Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., ... & Cohen, M. (2004, August). "Interactive digital photomontage. In ACM Transactions on Graphics (ToG) (Vol. 23, No. 3, pp. 294-302). ACM.(2004)

[34] ROTHER, Carsten; KOLMOGOROV, Vladimir; BLAKE, Andrew. "Grabcut: Interactive foreground extraction using iterated graph cuts." In: ACM transactions on graphics (TOG). ACM, p. 309-314.(2004)

[35] Li, Y., Sun, J., Tang, C. K., & Shum, H. Y. "Lazy snapping." ACM Transactions on Graphics (ToG), 23(3), 303-308.(2004

[36] C. Zhang and Z. Zhang, "A survey of recent advances in face detection," (2010)

[37] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," TPAMI, vol. 32, no. 9, pp. 1627.1645, (2010)

[38] B. Yang, J. Yan, Z. Lei, and S. Z. Li, "Aggregate channel features for multi-view face detection," in IJCB, pp. 1-8, IEEE, (2014)

[39] P. Viola and M. J. Jones, "Robust real-time face detection," IJCV, vol. 57, no. 2, pp. 137.154, (2004)

[40] J. Li, T. Wang, and Y. Zhang, "Face detection using surf cascade," in Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, pp. 2183-2190, IEEE, (2011)

[41] V. Kazemi and S. Josephine, "One millisecond face alignment with an ensemble of regression trees," in CVPR, pp. 1867-1874, IEEE Computer Society, (2014)

[42] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, "Face detection based on multi-block lbp representation," in International Conference on Biometrics, pp. 11-18, Springer, (2007)

[43] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5325-5334, (2015)

[44] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in NIPS, pp. 91.99, (2015)

[45] Liu, Wei, et al. "Ssd: Single shot multibox detector." European conference on computer vision. Springer, Cham(2016)

[46] E. Borenstein, E. Sharon and S. Ullman, "Combining Top-Down and Bottom-Up Segmentation." 2004 Conference

on Computer Vision and Pattern Recognition Workshop, Washington, DC, USA, pp. 46-46.(2004)

[47] Lienhart, R., & Maydt, J. "An extended set of haar-like features for rapid object detection." In: Image Processing. Proceedings. International Conference on IEEE(2002)

[48] Couprie, C., Grady, L., Najman, L., Talbot, H. "Power watersheds: A new image segmentation framework extending graph cuts, random walker and optimal spanning forest." In Computer Vision. In: 12th IEEE International Conference(2009)

[49] T. Mitsunaga, T. Yokoyama, and T. Totsuka, "Autokey: Human Assisted Key Extraction," Proceedings of ACM SIGGRAPH, pages 265-272, (1995)

감사의 글

샬롬 엘로이 에벤에셀

먼저 모든 영광을 하나님께 돌립니다. 그리고 하나님의 영광이 미물인 저로 인해 조금이라도 이 땅에 반사되기를 소망하면서, 14년 6개월간의 여정을 마무리하며 고마웠던 분들에게 감사의 인사를 드립니다. 20여년간의 오랜 시간동안 제자로 키워주신 우리 슈퍼바이저 지도 교수님 김영봉 교수님께 감사의 마음을 전합니다. 그리고 항상 극한 환경속에서도 옆에서 궂은 일도 마다하지 않고 지금껏 묵묵히 함께 생활하면서 많은 일을 도와준 권오석 박사에게도 미안한 마음과 함께 고맙단 말을 전합니다. 그리고 지금은 석사를 졸업하고 열심히 미래를 위해 준비중인 건국이에게도 좋은 후배가 되어줘서 고맙다는 말을 전합니다. 그리고 하나뿐인 박사과정 동기이자 친구인 착한 가림이, 먼저 졸업해서 미안한 마음과 긴 시간동안 챙겨줘서 고맙다는 말을 전합니다. 그리고 언제나 걱정스런 마음으로 챙겨주는 기진 누나 정말 고맙습니다. 그리고 수도 누나도 항상 친동생처럼 챙겨줘서 고맙습니다. 그리고 상화형님, 항상 어른스럽게 경험을 통한 삶의 지혜를 잘 가르쳐 주고, 때로는 편안한 친구처럼 부족한 동생을 잘 보살펴줘서 너무 고맙습니다. 그리고 지금 함께 일하면서 옆에서 저를 도와주시고, 챙겨주시는 우리 박경미 선생님 감사합니다. 그리고 지금은 서울에서 열심히 일하고 있는 연구실 석사 후배이자 내 동생 준영이에게 항상 진심으로 생각해줘서 고맙다는 말을 전합니다. 그리고 항상 남을 먼저 생각하고 배려하는 멋쟁이 조민석 대표님, 바쁜 와중에 항상 이 형아까지 전화로 챙기고 생각해줘서 고맙습니다. 그리고 지금 바로 옆에 있는 우리 부산 사나이 친구들, 먼저 나에겐 드웨인 존스보다 멋진 친구, 마동석보다 팔뚝이 굵지만 마음은 너무 착한 우리 회장님 민환(민규로 개명했어요)이 항상 지지해주고 살뜰히 챙겨줘서 고마운 마음입니다. 그리고 항상 새로운 일에 도전하고, 변화를 통해 점점 더 멋지게 변신해 가는 우리 구서동 맨하탄 휘트니스의 관장님, 내 친구 원석이 항상 걱정하고 챙겨줘서 고맙다는 말을 전합니다. 그리고 아직도 현역 선수 시절을 몸을 유지하며 무엇보다 내 건강을 생각해주는 몸짱이면

서, 심성이 너무 착한 우리 친구 강남구 코치님 고맙습니다. 그리고 나에게만 표현이 거칠지만 본성이 착하고 여린 내 친구 엄창준, 있는그대로의 진심으로 편하게 대해줘서 항상 고맙고, 감사하다는 말을 전합니다. 그리고 요즘 너무 바빠서 챙기지도 못하고 가끔씩 안부만 묻지만, 항상 마음으로 챙겨주고 생각하는 우리친구 김승구, 고마운 마음을 항상 갖고 있습니다. 그리고 지금은 서울에서 육아와 일로 바쁘게 지내는 우리 프로그래머 친구 최진우, 그리고 프로그래머에서 금융맨으로 변신한 태균이, 그리고 부동산 투자의 고수가 된 상국이 이들에겐 항상 내가 빚진마음과 함께 고마운 마음입니다. 그리고 항상 세상의 궂은 프로그램 개발일을 다하면서, 이 동생을 챙겨주는 원경이형, 그리고 지금 너무 멋지게 살고 있어서 뉴스로 소식을 듣는 창율이형 내 마음 속에 항상 고마움으로 남아있습니다. 그리고 제 영혼의 고향인 부산성동교회 교우분들께도 항상 미안한 마음과 고마움 마음을 전합니다. 그리고 지금 제가 파트타임으로 일하고 있는 부산IT직업학교의 선생님들께도 항상 스케줄도 배려해주시고 챙겨주셔서 고맙다는 말을 전합니다.

마지막으로 이 부족하고 모자란 다 큰 동생을 믿고 지지해주시고, 경제적 지원까지 해주신 우리 현우 형님 쑥스럽지만 고맙다는 말을 전합니다. 그리고 아들이 어떤 일을 하던간에, 언제나 믿음으로 사랑으로 기도로 챙겨주고 보살펴 주신 우리 어머니 김선애 여사님 미안하고, 사랑하고 고맙습니다. 당신의 헌신과 사랑이 있었기에 이 모든 일이 있었다고 생각합니다. 다시 한 번 어머님께 사랑한다는 말을 전하며, 부끄럽지만 자랑스럽고, 미안하지만 고마운 이 마음을 담아 이 논문을 우리 어머니 김선애 여사님께 바칩니다.

- 윤 요 섭