

게임 사용자 얼굴 검출을 위한 3D 공간 영상 추출 기법

유채곤, *정창성, *황치정
대덕대학 컴퓨터정보통신계열 컴퓨터게임전공
* 충남대학교 영상처리 연구실

Image Extraction Method in 3D Space for Game Player's Face Detection

Chae-Gon Yoo, *Chang-Sung Jung, *Chi-Jung Hwang
Computer Game Major, Daeduk College
*Image Processing Lab. Chungnam National University

요약

본 논문에서는 복잡한 랜덤 배경 하에서 위치하고 있는 게임 플레이어의 얼굴 영상을 스테레오 매칭을 이용하여 배경과 분리하여 추출할 수 있는 방법에 대하여 기술한다. 사람과의 상호 작용이 필요한 게임일수록 사람의 동작이나 각 부위에 대한 인식이 필요하다. 이 방법은 게임 이외에도 보안 시스템, 의류 시뮬레이션, 3D 모델링 그리고 로보틱스와 같은 분야에 적용될 수 있다. 스테레오 매칭에 관해서는 많은 연구가 있어왔으며, 기본적으로 영역기반 방법과 특징기반 방법으로 분류될 수 있다. 본 논문의 제안 방법에서는 영역기반 방법으로 처리를 시작하고, 다단계 크기의 윈도우를 적용하여 물체의 경계선을 찾는 작업을 진행한다. 각 윈도우 크기에 대하여 유사성 커브가 생성되며, 이 값은 물체의 경계선을 판별하는 특징으로 사용된다. 전단계에서 생성된 코어스(coarse) 영역은 유사성 커브 방식에 의하여 머지 작업을 거치며, 최종적으로 대상 물체의 영상을 추출하게 된다.

1. 개요

스테레오 매칭 기법은 컴퓨터 비전 분야에서 자주 사용되는 방법이다. 스테레오 페어 영상간 비교를 위하여, 영역기반 방법에서는 영역 기반 비교 방법이 사용되며, 특징 기반 방법에서는 에지나 코너와 같은 특징들이 참조 포인트로서 사용된다. 두 가지 방법은 여전히 열악한 텍스처 데이터, 반복 패턴, 윈도우의 크기, 폐색(occlusion)과 같은 현상에 대해서는 문제점을 가지고 있다. 이러한 문제점들 때문에, 빌딩이나 탑과 같이 상대적으로 규칙적인 특징을 가지고 있는 대상에 대해서는 스테레오 매칭을 이용한 3D 재구성 등의 기법이 용이하게 적용되지만, 나무나 사람과 같이 랜덤한 물체에 대해서는 많은 어려움을 가지고 있다.

본 논문에서는, 영역을 기반으로 하여 스테레오 영상의 두 영상간에 깊이(depth) 분포를 구하는 coarse 영상 매칭

을 수행한 후 목표 물체의 영역을 구하기 위하여 머지 작업을 수행한다. 다단계 크기의 윈도우를 적용하여 물체의 경계선 부분의 유사성을 측정하기 위한 유사성 커브 방법이 소개된다. 유사성 커브는 목표 물체의 경계선 특성에 대한 정보를 가진다. 본 논문의 제안 방법은 블루 스크린과 같은 특정한 상황 설정을 하지 않고도 보안 시스템, 사람의 얼굴 영상 취득을 통한 의류 시뮬레이션, 3D 모델링, 로보틱스와 같은 컴퓨터 비전 분야에 적용될 수 있다.

2장에서는 관련된 연구에 대한 리뷰를 하고, 3장에서는 제안 방법이 기술되며, 4장에서는 실험결과가 도시된다. 마지막으로 5장에서는 결론과 향후 연구가 설명된다.

2. 기존 연구

대부분의 기존 연구는 크게 특징 기반 방법과 영역 기반

방법의 범주로 구분된다. 영역 기반 방법에서는 비연결적 공간을 찾기 위하여, 폐색(occlusion)을 검색하기 위한 방법이 시도되었다[1]. 또 다른 시도에서는 스테레오 영상을 디지털 엘리베이션과 디지털 터레인의 두 가지 모델로 분할을 시도하였다[2]. 미지의 정적인 장면의 3D 재구성을 수행하기 위하여 단일 이동 카메라를 사용하여 비디오rate으로 정확한 3D 재구성을 시도하기도 하였다[3]. 물체 인식을 위하여 광도적 인베리언트와 기하학적인 인베리언트를 사용하는 연구도 꽤되었다[4].

3D 모델링과 뷰포인트 합성 방법도 스테레오 매칭의 방법으로 시도되었다[5]. 스테레오 매칭의 정확한 영상 취득 환경을 위한 캘리브레이션에 대한 연구로서, 특징을 추출하지 않고도 캘리브레이션을 처리할 수 있는 방법도 연구가 되었다[6]. 통계적인 방법 중 Maximum Likelihood 방법을 적용하여 영상의 소속 영역 결정을 위한 연구에서는 코스트 함수를 사용하였다[7]. 스테레오 매칭에서 특징을 계층적으로 구분한 예로 V. Venkateswar는 특징의 종류를 라인, 점, 에지 그리고 면으로 계층적으로 분류하고, 최고 레벨의 특징으로부터 매칭을 시작하여 가장 낮은 레벨로 진행하는 방법을 소개하였다[8]. Sebastien Roy는 Epipolar 라인을 기하를 사용하지 않고, 스테레오 매칭의 요소를 그래프의 Maximum Flow 문제로 변환하여 해결하고자 하였다[9].

본 제안 방법에서는 깊이 계산과 유사성 커브 방식을 사용하여 경계선 영역 측정을 시도하는 방식이며, 기존의 연구와는 달리 실행 시간을 줄이기 위하여 2D 도메인에서 깊이에 따른 물체의 분포와 배경과의 비교에 초점을 두었다.

3. 목표 물체 영상 추출 방법

프로세싱의 전반부에서는 처리속도를 위하여 quad-tree를 이용한 코어스(coarse) 영역 매칭을 수행한다. 코어스 매칭에서 처리된 영역에 대한 중간 깊이 계산을 수행한다. 중간 깊이 정보는 특정 깊이 threshold에 기반하여 영역 머지에 이용된다. 다음 단계에서는 fine tuning 단계에서 선정된 영역에 대하여 정확한 경계선 검출을 시도한다.

그림 3-1에서는 본 제안 방법이 적용될 시스템을 도시하고 있다. 두 대의 카메라를 사용하여, 사용자의 얼굴 부분을 불규칙적이고 복잡한 배경으로부터 분리하여 추출하는 것

이 본 시스템의 목적이다. 본 시스템이 게임기와 같은 실제 응용 시스템에 사용되기 위해서는 실시간에 작업이 처리되어야 하므로, 스테레오 매칭의 수행시간에 대한 제한도 본 논문의 전제가 된다.

경계선에 위치한 픽셀의 판단을 위하여, 본 제안 방법에서는 윈도우의 크기를 변화하여 해당 픽셀에 대하여 유사

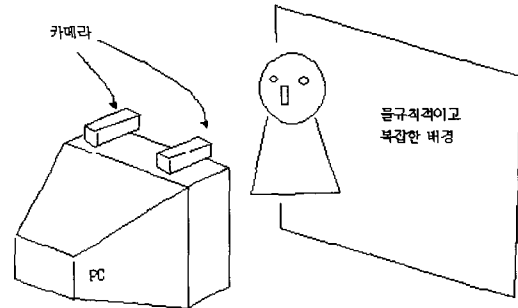


그림 3-1. 시스템 적용도

성 커브를 생성한다. 경계선의 일부로 간주된 픽셀들은 물체 영상 추출 단계로 전달된다. 물체의 배경이 블루스크린과 같이 특정 환경이 아닌 랜덤 환경인 경우에도, 제안 방법은 영향을 받지 않는 특성을 가지고 있다. 그림 3-2는 전체 시스템의 흐름을 도시한다.

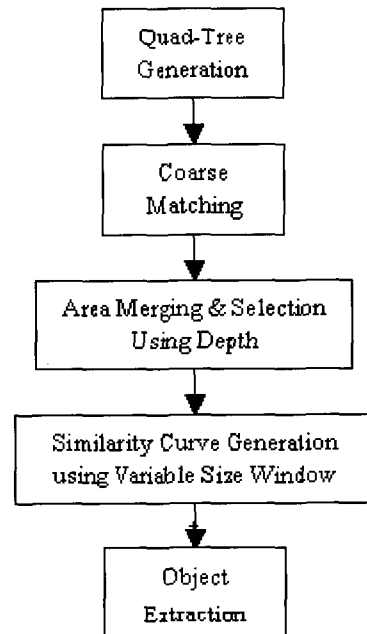


그림 3-2. 시스템 구성도

3.1. 두 영상간 밝기 정규화

영상은 날씨, 촬영 시각과 같은 획득 당시의 환경에 의하여 전체 화소의 밝기값에 차이가 발생할 수 있으며, 밝기값의 차이에 의하여 히스토그램의 패턴도 영향을 받게 된다. 광원에 의한 화소의 밝기값이 달라질 경우 히스토그램의 분포 위치만 이동하게 된다. 히스토그램의 모양이 유사하더라도 분포 위치가 변하면 상관 계수나 변이 계수 측정시 많은 오차가 발생되므로, 정규화 과정을 통하여 히스토그램 분포의 위치를 동일하도록 만들어줄 필요가 있다.

매칭의 대상이 되는 두 비교 영상간에 밝기값 정규화를 시킬 때 어떤 쪽의 영상을 기준으로 정규화를 하더라도 무방하다. 두 영상의 밝기값 분포에 대한 정규화 과정은 식 (1)과 같다.

$$f(x, y) = m_A + (f(x, y) - m_B) * \frac{\sigma_A}{\sigma_B} \quad (1)$$

where

A : 영상 A

B : 영상 B

m_A, m_B : A, B 의 밝기값의 평균

σ_A, σ_B : A, B 의 밝기값의 표준편차

3.2. 영역 기반 매칭

영역 기반 매칭 단계에서는 스테레오 영상 비교의 속도를 높이기 위하여 quad-tree 피라미드가 사용된다. 그림 3-3에서는 quad-tree 피라미드 구조를 보여주고 있다. Quad-tree 의 레벨은 요구되는 실행 시간에 따라서 조정될 수 있다. 각 영상의 서브 블록은 그레이 레벨 밝기 값의 상관계수를 이용하여 유사도가 측정된다.

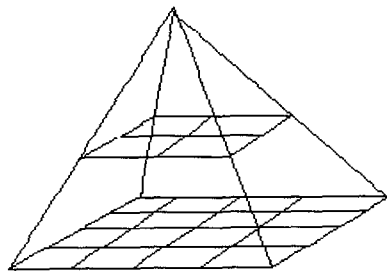


그림 3-3. Quad-tree 피라미드

3.3. 유사성 커브

이제, 영상 추출 대상이 되는 물체의 영역의 윤곽이 결정된 상태이므로, 물체의 경계선 영역을 판별하여 물체 영상의 내부의 영역들은 머지하는 과정이 필요하다. 물체의 내부에 있는 픽셀은 큰 크기의 비교 윈도우를 사용하더라도, 두 개의 스테레오 영상의 해당 위치 지점에서 유사한 유사성 커브 값을 가지는 특성을 가지고 있다. 반면, 물체의 경계선에 가깝게 위치한 픽셀일수록 비교 윈도우의 크기가 threshold에 접근할수록 급격하게 유사성 값이 감소하는 특성을 가지고 있다.

$$C_{xyw} = \sqrt{\int_{x-i}^{x+i} \int_{y-i}^{y+i} (li(m, n) - ri(m, n + w))^2 dmdn} \quad (2)$$

where

x, y : 픽셀의 좌표

i : 비교 윈도우 크기

li : 좌측 영상 픽셀의 밝기 값

ri : 우측 영상 픽셀의 밝기 값

w : 두 개의 영상 내에서 목표 물체의

x 좌표 변이값

식 (2)의 w 는 영역 기반 매칭 단계에서 물체의 거리에 반하여 결정된다. 만약 w 를 직접 계산하지 않고, 거리에 따른 매핑 테이블을 사용하여 미리 기억시켜 놓는 경우에는 처리 시간을 많이 단축시킬 수 있으며, 보다 정확한 결과를 생성할 수 있다. 시스템의 캘리브레이션 단계에서 초점 거리, 카메라의 베이스 라인, 물체의 거리를 입력으로 받아 목표 물체에 대한 w 를 미리 계산해 놓을 수 있다. 사각형 윈도우는 물체의 모양에 따라서 다른 결과를 생성할 수 있기 때문에, 유사성 커브 생성을 위해서는 원형 비교 윈도우를 사용한다.

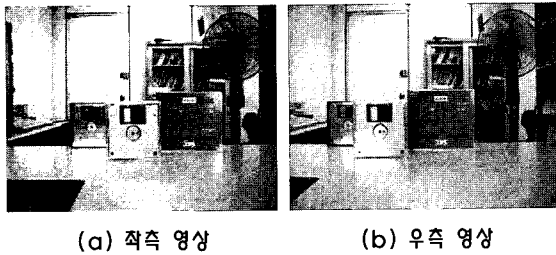
3.4. 물체 영상 추출

본 단계에서는, 대상물체 영상이 백그라운드로부터 분리된다. 물체의 내부로 간주되어 머지된 영역은 목표 물체의 영상 영역으로 간주되며, 경계선에 가까운 부분이라도 상대적으로 높은 유사성 커브 값을 가지고 있는 경우에는 목표 물체의 영상 영역으로 간주된다. 이 기준에 의하여 우리는 블루스크린과 같은 특정 환경을 이용하지 않고도 랜덤 백그라운드로부터 목표 물체의 영상을 분리할 수 있다.

4. 실험 결과

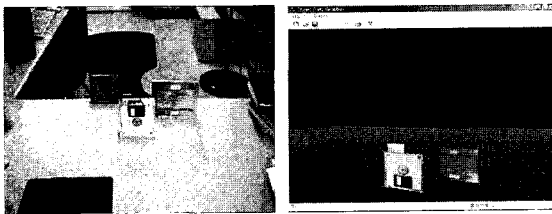
실험을 위하여 사용된 시스템은 펜티엄 III 700Mhz cpu와 256M 의 메모리를 사용하였다. 스테레오 영상의 크기는 640 x 480 해상도를 가진다. 실험에는 물체와 사람얼굴을 대상으로 한 두 개의 실험 영상 세트를 사용하였다. 그림 4-1과 그림 4-2의 영상 (a)와 (b)는 스테레오 영상의 좌측과 우측의 영상이다. 그림 4-1에서 영상 추출 목표 물체는 가장 앞에 놓여진 박스이며, 그림 4-2에서는 사람 얼굴이 영상 추출 목표 물체이다.

영역 기반 매칭 단계에서 물체의 대략적인 윤곽을 얻을 수 있으며, 이 단계에서 노이즈나 좌우측 영상간의 밝기 차이에 의해서 잘못 계산된 에러가 발생할 수도 있다. 이런 노이즈 들은 fine 튜닝단계에서 주변 픽셀들과의 상관관계를 비교하여 제거된다. 경계선 영역의 각 픽셀에 대해서는 유사성 커브를 생성하여, 물체의 영역으로의 간주 정도를 계산한다.



(a) 좌측 영상

(b) 우측 영상



(c) 오버뷰 영상

(d) 이미지 추출 영상

그림 4-1. 물체 실험 영상

영역 기반 매칭과 유사성 커브에 기반한 방법을 사용하여, 첫 번째 실험에서는 그림 4-1의 (d)와 같은 결과 영상을 생성하였다. 첫 번째 시험의 경우는 대상 물체의 모양이 사각형과 같은 규칙적인 모양을 하고 있는 경우이다. 목표 영상은 랜덤 백그라운드로부터 분리된 것을 볼 수 있다. 그림 4-1이 (d)는 생성된 결과를 계산된 깊이에 따라서 3D 공간



(a) 좌측 영상

(b) 우측 영상



(c) 영역 검출 영상



(d) 얼굴 검출 영상

그림 4-2. 사람 실험 영상

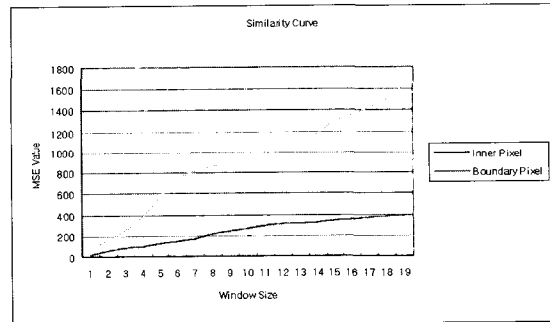


그림 4-3. 유사성 커브

에 위치시키고 네비게이션하면서 생성한 화면이며 영상간 분리 결과를 확인할 수 있다.

두 번째 시험의 경우는 그림 4-2에 나타나 있으며, 사람의 얼굴의 모양이 불규칙한 경계선을 가지고 있다는 것만 제외하고는 첫 번째 실험과 내용은 동일하다. 그림 4-2의 (c)는 좌측과 우측 스테레오 영상인 그림 4-2의 (a)와 (b)를 자동 검색하여 얼굴의 위치를 일차로 검색한 결과이다. 그림 4-2의 (d)는 일차 검색 결과에 대하여, 경계선 측정을 시도하여 얼굴 영상을 배경으로부터 추출한 최종 결과이다. 처리 시간은 앞에서 설명한 프로세서와 메모리 환경에서 최종 결과를 생성하기까지 1초 이내로 처리되므로, 실용적인 적용의 경우에도 문제가 없다고 판단된다.

그림 4-3의 그래프는 물체의 내부에 있는 픽셀과 경계선

부분에 있는 픽셀의 유사성 커브 변화 정도를 도시하고 있다. 이 그래프는 MSE(Minimum Square Error) 값이므로, 값이 클수록 유사성이 떨어짐을 의미한다.

경계선 부근의 픽셀일 경우 내부에 있는 픽셀의 경우보다 에러의 증가비율이 급격하게 높은 것을 볼 수 있다.

5. 결론 및 향후 연구

실험 결과에서는 제안 방법이 목표영상이 랜덤 백그라운드로부터 분리되었음을 보이고 있다. 본 제안 방법은 게임 시스템에서 사용자 인식, 물체 인식, 의류 시뮬레이션, 보안 시스템, 로봇틱스와 같은 컴퓨터 비전 분야에 적용될 수 있다. 실험 결과에서는 블루스크린과 같은 특정 배경을 설정하지 않고도 원하는 물체 혹은 사람의 얼굴의 위치와 영상의 추출에 성공하였다. 사람 얼굴의 경우 경계선에서 머리카락의 모양과 색상이 불규칙하여 경계선 결정에서 약간의 오차가 발생하였지만, 현재 결과로서 취득된 사람얼굴 영역 영상의 경우로도 다양한 응용이 가능하다.

향후에는, 영역 기반 방법과 특징 기반 방법을 함께 사용하여 정확도를 높이는 방법에 대한 연구가 필요하며, 본 실험에서 사람의 경우에 발생한 불규칙한 모양과 색상을 가진 머리카락 부분의 경계선 처리 문제를 해결할 수 있는 방안에 대한 연구가 필요하다.

참고 문헌

[1] Aaron F. Bobick, "large Occlusion Stereo," International Journal of Computer Vision, 33(3), pp 181-200, 1999

[2] C. Baillard, H. Maitre, "3-D Reconstruction of Urban Scenes from Aerial Stereo Imagery: A Focusing Strategy," Computer Vision and Image Understanding Vol. 76, No. 3, December, pp. 244-258, 1999

[3] Eric Marchand, Francois Chaumette, "An Autonomous Active Vision System for Complete and Accurate 3D Scene Reconstruction," International Journal of Computer Vision 32(3), 171-194, 1999

[4] Kenji Nagao, "Using Photometric Invariants for 3D Object Recognition," Computer Vision and Image

Understanding, Vol. 71, No. 1, July, pp74-93, 1998

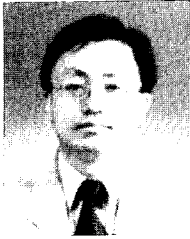
[5] Ebroul Izquierdo M. and Silko Kruse, "Image Analysis for 3D Modeling, Rendering, and Virtual View Generation," Computer Vision and Image Understanding Vol. 71, No. 2, August, pp. 231-253, 1998

[6] Luc Robert. "Camera Calibration without Feature Extraction" , Computer Vision and Image Understanding, Vol 63, No. 2, pp. 314-325, 1996

[7] Bruce M. Maggs. "A Maximum Likelihood Stereo Algorithm" , Computer Vision and Image Understanding, Vol 63, No. 3, pp. 542-567, 1996

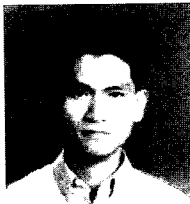
[8] V. Venkateswar and R. Chellappa. "Hierarchical Stereo and Motion Correspondence Using Feature Groupings" , International Journal of Computer Vision, Vol 15, pp. 245-269, 1995

[9] Sebastien Roy. "Stereo Without Epipolar Lines : A Maximum-Flow Formulation" , International Journal of Computer Vision, Vol 34, pp. 147-161,



유채곤

1992 충남대학교 전산학과 (이학사)
 1997 충남대학교 전산학과 (이학석사)
 1998 South Dakota 주립대 전산과 (이학석사)
 2001 충남대학교 컴퓨터과학과 (박사수료)
 2000년-현재 대덕대학 컴퓨터게임전공 주임교수
 관심분야: 영상처리, 가상현실, 컴퓨터비전



정창성

2000 한밭대학교 컴퓨터공학과 (공학사)
 2000-현재 충남대학교 컴퓨터과 학과 석사과정
 관심분야: 영상처리, 컴퓨터비전, 멀티미디어



황치정

1975 서강대학교 수학 (이학사)
 1979 서강대학교 수학 (이학석사)
 1980 SUNY at Albany 수학(석사)
 1983 Univ. of Connecticut 수학 (박사 수료)
 1985 Univ. of Connecticut 전산학 (이학 석사)
 1987 Univ. of Connecticut 전산학 (이학 박사)
 1997-1999 충남대학교전자계산소장
 1988-현재 충남대학교 컴퓨터과학과 교수
 2001-현재 Las Vegas 주립대 교환교수
 관심분야: 영상처리, 컴퓨터비전