

## 구조 광 영상을 이용한 3차원 얼굴 복원

### 3D Faces Reconstruction Using Structured Light Images

이득용, 오일석  
전북대학교 전자정보공학부 컴퓨터공학전공

Duk-Ryong Lee, Il-Seok Oh  
Computer Engineering Major,  
Division of Electronics and Information Engineering,  
Chonbuk National University

#### 요약

본 논문은 구조 광 영상을 이용하여 3차원 얼굴을 복원하는 방법을 제안한다. 프로젝터와 카메라의 시선벡터가 평행하다고 가정한다. 먼저 격자 형태의 구조 광을 배경에 투영하고 참조 구조 광 영상을 획득한다. 참조 구조 광 영상을 사용하여 카메라와 프로젝터를 보정한다. 이후 사람의 얼굴에 동일한 형태의 구조 광을 투영하고 얼굴 구조 광 영상을 획득한다. 획득한 두 종류의 구조 광 영상에서 추출된 특징 점들의 위치 변화를 측정하고 측정된 변화량으로 3차원 얼굴을 복원 한다. 실험 결과를 보면, 이런 간단한 장치를 통해 사람이 인식할 수 있을 정도의 3차원 얼굴 영상을 얻을 수 있음을 알 수 있다.

#### Abstract

This paper proposes a method to reconstruct the 3-D face using structured light image. First of all, we suppose that each sight vector of a projector and camera are parallel. We project the structured light in the shape of lattice on the background to acquire the reference-structured light image. This image is used to calibrate the projector and camera. Since then, we acquire the face-structured light image which is projected the same structured light on the face. These two structured light images are used to reconstruct the 3-D face through the variation which is measured from the positional difference of feature vectors. In our experiment result, we could reconstruct the 3-D face image as recognize through these simple devices.

## 1. 서론

### 1.1 연구배경

과거 3차원 콘텐츠는 의복 전자상거래 시스템, 온라인 게임, 애니메이션과 같은 한정된 분야에 활용되었다. 하지만 인터넷이 대중화됨에 따라 3차원 콘텐츠의 활용 분야는 특정 영역에서 개인 영역으로 점차 확대되고 있다. 즉 개인의 성향을 반영한 3차원 콘텐츠의 개발이 필요하게 된 것이다. 개인의 성향을 반영하는 가장 효과적인 방법 중의 하나는 3차원 얼굴을 생성하는 것이다. 하지만 개인이 3차원 얼굴 영상을 생성하는 것은 매우 어렵다 [1, 2].

3차원 얼굴 영상을 생성하는 대표적인 방법으로는 3차원 입체 스캐너 방법, 스테레오 카메라 방법, 구조 광(Structured Light) 방법 등이 있다. 3차원 입체 스캐너를 사용하면 매우 정교한 영상을 얻을 수 있다 [3]. 하지만 고가의 장비를 구축해야 하며, 입체 스캐너가 설치되지 않은 장소에서는 사용할 수 없다. 스테레오 카메라 방법은 현재 많은 연구가 이루어지고 있으며 비교적 높은 성능을 보이고 있다 [4, 5]. 하지만 다수의

카메라를 사용하기 때문에, 개인이 이 방법을 통해 3차원 얼굴 영상을 생성하기에는 무리가 있다. 구조 광 방법은 스테레오 카메라 방법에서 사용하는 카메라 중 하나를 프로젝터로 대체한 방법으로 프로젝터를 사용하여 얼굴에 구조 광을 투영하고 투영된 구조 광을 특징으로 사용하여 3차원 영상을 복원하는 방법이다 [6, 7].

구조 광 방법을 사용하기 위해서는 구조 광을 투영하기 위한 별도의 하드웨어가 필요한데 과거에는 매우 높은 가격이었다. 하지만 최근에 프로젝터가 널리 보급되어 구조 광 영상을 매우 쉽게 얻을 수 있게 되었다. 즉 별도의 하드웨어를 설계하지 않고 학교, 실험실 및 가정에서 손쉽게 구조 광 영상을 얻을 수 있게 된 것이다.

### 1.2 연구 내용

본 논문에서는 구조 광 영상에서 3차원 얼굴 영상을 복원하는 방법을 제안한다. 먼저 프로젝터를 사용해 구조 광을 배경에 투영하고 참조 구조 광 영상을 얻는다. 참조 구조 광 영상은 특징 점의 위치 변화를 측정하기 위해 사용된다. 이후 동일

한 형태의 구조 광을 얼굴에 투영하고 얼굴 구조 광 영상을 획득한다. 얼굴 구조 광 영상은 얼굴의 영향으로 특징 점의 위치가 변하게 된다. 따라서 두 구조 광 영상의 특징 점의 위치 변화를 측정하여 3차원 깊이 정보로 변환 한다.

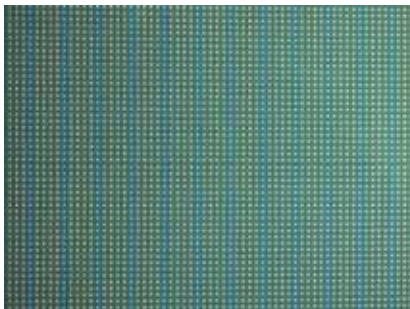
논문의 2장에서는 3차원 얼굴 복원 알고리즘을 설명하고 3장에서 제안하는 알고리즘을 구현하고 실험 결과를 제시한다. 4장에서는 제안하는 방법의 장점과 결론을 기술한다.

## 2. 얼굴 복원 알고리즘

### 2.1 구조 광

3차원 얼굴 영상을 복원하기 위해서는 투영하는 구조 광의 위치, 크기, 형태 등을 결정해야 한다. 제안하는 방법은 투영된 특징 점의 위치 변화를 기반으로 깊이를 계산하기 때문에 선분이 아닌 점들을 투영 한다. 특징 점들은 상하좌우 동일한 간격을 가지는 메쉬 형태이다. [그림 1]은 프로젝터에서 배경에 투영한 구조 광을 카메라로 찍은 영상이다. 배경 구조 광 영상은 보정 과정과 깊이 계산 과정에서 참조 영상으로 사용된다.

[그림 2]는 얼굴 구조 광 영상이다. 배경 구조 광 영상과 비교해 보면, 얼굴에 투영된 특징 점들은 위치의 변화가 있으나 배경에 투영된 특징 점들은 위치 변화가 없음을 확인할 수 있다. 얼굴에 투영된 특징 점과 배경에 투영된 특징 점의 위치 변화를 측정하려면 두 특징 점이 동일하다는 전제를 만족해야 한다. 서로 매치되는 특징 점의 쌍은 2차원 메쉬 추출 방법으로 얻는다 [8].



▶▶ 그림 1. 배경 구조 광 영상



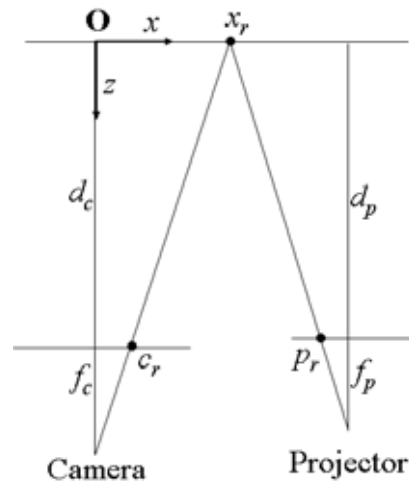
▶▶ 그림 2. 얼굴 구조 광 영상

### 2.2 카메라와 프로젝터의 기하 모델

카메라와 프로젝터의 기하 모델에서는 [그림 3]과 같이 카메라와 프로젝터의 시선 벡터(Line of Sight)가 평행 하다고 가정한다. 카메라와 프로젝터를 평행으로 가정 할 경우 보정(Calibration) 과정이 매우 간단해진다. [그림 3]은 프로젝터가 점을 투영하고 카메라가 그 점을 얻는 과정을 보여준다. 프로젝터 좌표계의 점  $p_r$ 을 세계 좌표계(O)에 투영하고, O의 점  $x_r$ 은 카메라 좌표계(Camera coordinate:C)의 점  $c_r$ 이 된다.

[표 1] 기하 모델에 사용된 파라미터

|       |                        |
|-------|------------------------|
| $f_c$ | 카메라의 초점 거리             |
| $f_p$ | 프로젝터의 초점 거리            |
| $d_c$ | 세계좌표계 x-y평면에서 카메라까지 거리 |
| $d_p$ | 프로젝터의 렌즈와 O의 x축과의 거리   |
| $d$   | 카메라와 프로젝터의 거리          |
| $p_r$ | 프로젝터 좌표계의 점            |
| $x_r$ | 점 $p_r$ 이 O의 x축에 투영된 점 |
| $c_r$ | 점 $x_r$ 이 카메라에 투영된 점   |



▶▶ 그림 3. 카메라와 프로젝터의 기하 모델 (x,z평면)

[그림 3]의 기하 모델은 다음 수식을 만족한다.

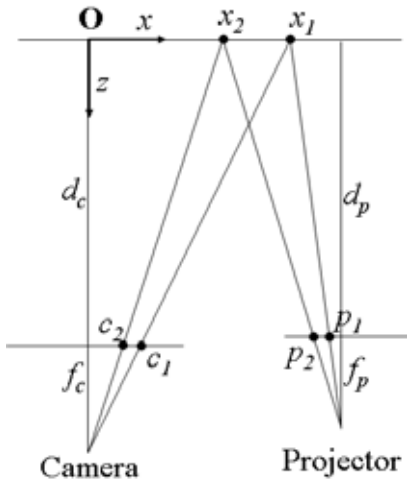
$$x_r : c_r = f_c + d_c : f_c \tag{1}$$

$$d - x_r : p_r = f_p + d_p : f_p \tag{2}$$

### 2.3 보정

위 기하 모델은 하드웨어 환경을 구축하는 과정에서  $c_r$ ,  $d_c$ ,  $d_p$ 는 측정이 가능하지만  $f_c$ 와  $f_p$ 는 측정이 불가능 하다. 3차원 정보를 얻기 위해서는  $f_c$ 와  $f_p$ 를 알고 있어야 한다.  $f_c$ 와  $f_p$ 를 계산하기 위해  $p_1$ ,  $p_2$ 를 투영한다. [그림 4]의 점들은 다음 수식

을 만족 한다.



▶▶ 그림 4. 카메라 보정 과정 (x,z평면)

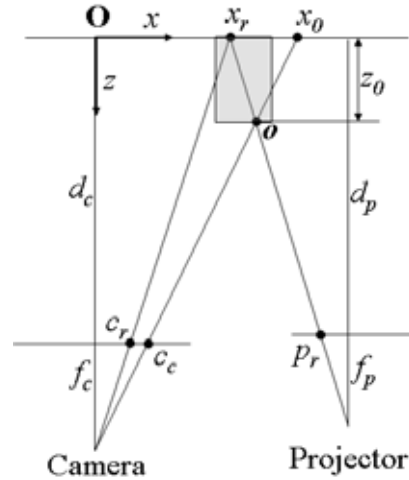
$$x_1 - x_2 : c_1 - c_2 = f_c + d_c : f_c \tag{3}$$

$$x_1 - x_2 : p_2 - p_1 = f_p + d_p : f_p \tag{4}$$

위 수식에서  $x_1 - x_2, c_1 - c_2, p_2 - p_1$ 는 측정이 가능하다. 이때  $x_1 - x_2, p_2 - p_1$ 은 실험 설정 단계에서 설정 해 주어 야 하며  $c_1 - c_2$ 은 참조 구조 광 영상에서 얻어진다. [그림 1]과 같이 배경에 투영된 특징 점들을 대상으로 위 수식을 적용하 면  $f_c$ 와  $f_p$ 를 계산 할 수 있다.

### 2.4 깊이 계산

이 과정에서는 배경 구조 광 영상의 특징 점  $c_r$ 과 얼굴 구조 광 영상의 특징 점  $c_c$ 의 거리 차이를 이용하여  $o$ 의 깊이( $z_0$ )를 계산한다. [그림 5]의  $o_r$ 은 사람의 얼굴에 의해 발생한 특징 점의 위치 변화를 계산하기 위한 참조 특징 점이며  $o$ 는 사람의 얼굴에 투영된 특징점이다. 사람의 얼굴에 투영된 특징 점의 위치 변화는 [그림 2]에서 확인 할 수 있다. 사람의 얼굴에 투영된 특징 점  $o$ 가 카메라에 투영된 점을  $c_c$ 라 하면, 각 점들의 관계는 아래 수식과 같다.



▶▶ 그림 5. 깊이 계산 모델 (x,z평면)

$$x_0 = x_r + z_0 \left( \frac{p_r}{f_p} + \frac{c_c}{f_c} \right) \tag{5}$$

$$f_c : f_c + d_c = c_c : x_0 \tag{6}$$

식 (5)를 식 (6)에 대입하여 정리하면

$$z_0 = \frac{(c_c - c_r)(f_c + d_c)f_p}{(p_r f_c + c_c f_p)} \tag{7}$$

$c_c$ 와  $c_r$ 은 2차원 메쉬를 추출함으로써 각 점마다 계산 할 수 있으며, 나머지 변수들은 이미 측정하거나 보정과정에서 계산 해 주었다. 따라서 [그림 2]의 모든 점을 대상으로 식 (7)을 적용하면 각 점에서의 깊이를 얻을 수 있다.

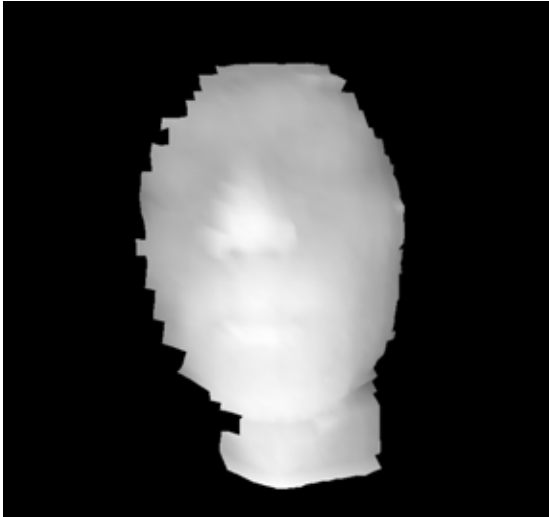
### 3. 실험 및 분석

제안하는 알고리즘의 성능을 평가하기 위하여 HVR-2000 카메라를 이용하여 1024x768의 해상도를 가진 구조 광 영상을 획득하였다. 프로젝터는 HITACHI CP-X430을 사용하였 다.

카메라와 프로젝터의 시선 벡터를 평행하게 하기 위해서는 하드웨어를 구축하지 않고 간단한 프로젝터의 보정 기능을 사용하였다. 프로젝터에서 투영한 특징 점은 62x48개의 점을 투영하였으며, 이때 약 35x40 개의 특징 점이 얼굴에 투영되었 다. 즉 3차원 얼굴은 35x40개의 점으로 구성되어 진다. 알고리 즘 적용 후 3차원 얼굴 영상은 OpenGL을 사용하여 렌더링 하 였다.

계산에 사용된 파라미터들은 다음과 같다.

- $d_c = 120$  cm
- $d_p = 120$  cm
- $x_1 - x_2 = 0.5$ cm
- $p_2 - p_1 = 8$  pixels



▶▶ 그림 6. 3차원 얼굴 영상

[그림 1]을 참조 구조 광 영상으로 [그림 2]를 얼굴 구조 광 영상으로 사용하여 제안하는 방법을 적용한 결과 [그림 6]과 같이 이목구비를 확인 할 수 있는 영상을 얻을 수 있었다. 하지만 특징 점을 추출 하는 과정에서 발생한 오류가 코 부분의 일그러짐 현상을 발생 시켰다. 즉 특징 점 추출 과정이 제안하는 알고리즘의 성능을 결정 했다고 할 수 있다. 또한  $x_1 - x_r$ 의 거리가 3차원 영상의 해상도를 결정하게 된다.

#### 4. 결 론

본 논문에서는 3차원 얼굴 영상을 복원하기 위한 프로젝터와 카메라의 기하 모델을 정의 하였다. 또한 기하 모델을 바탕으로 카메라와 프로젝터를 보정하고, 보정 정보를 이용하여 3차원 얼굴을 복원하기 위한 식을 유도하였다.

제안하는 방법은 구조 광 영상을 투영하기 위해 프로젝터를 사용하기 때문에 고가의 하드웨어를 구축하지 않고 3차원 얼굴 영상을 얻을 수 있었다. 즉 개인이 손쉽게 사용 할 수 있기 때문에 다양한 환경에서 다양한 사용자가 3차원 콘텐츠를 얻을 수 있음을 확인 하였다.

향후 연구로는 3차원 얼굴 영상의 해상도를 높이는 방법에 대한 연구가 필요하다.

#### ■ 참 고 문 헌 ■

- [1] 정일홍, 조세홍, "실시간 감정 표현 아바타의 설계 및 구현," 디지털콘텐츠학회 논문지, vol. 7, no. 4, pp.235-243, 2006.
- [2] 여영인, 서혜원, 원광연, "이미지 입력을 이용한 3차원 아바타의 데이터기반 생성기법," 정보과학회 한국컴퓨터종합학술대회 논문집, vol. 2005, no. 0, pp.682-684, 2005.
- [3] S. Marschner, B. Guenter, and S. Raghupathy, "Modeling and rendering for realistic facial animation," *Proc. Eurographics Workshop on Rendering*, 2000.
- [4] T. D. Williams, "Depth from camera motion in a real world scene," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-2, no. 3, pp. 511-516, 1980.
- [5] R. Jain, S. L. Bartlett, and N. O'Brien, "Motion stereo using egomotion complex logarithmic mapping," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 3, pp. 356-369, 1987.
- [6] Datong Chen, Wen Gao, and Xilin Chen, "A new approach of recovering 3-D shape from structure-lighting," *International Conference on Signal Processing*, Vol. 2, pp.839-842, 1996.
- [7] 도경민, 이쾌희, "동적 스테레오 카메라 모델링," 제어자동화시스템공학회 논문지, vol. 3, No. 3, 1997.
- [8] 이득용, 오일석, "3차원 얼굴 복원을 위한 구조 광 영상에서의 2차원 메쉬 추출," 한국콘텐츠학회 종합학술대회 논문집, vol. 5, no. 1, pp. 248-251, 2007.