

# 실세계 객체의 디퓨즈 텍스처 획득에 관한 연구

김강연<sup>1</sup>, 이재영<sup>2</sup>, 유재덕<sup>3</sup>, 이관행<sup>4</sup>  
광주과학기술원<sup>1 2 3 4</sup>  
{k2y00<sup>1</sup>, jylee81<sup>2</sup>, uranus<sup>3</sup>}@gist.ac.kr, lee<sup>4</sup>@kyebek.gist.ac.kr

## Study on the Diffuse Texture Acquisition of a Real Object

Kang Yeon Kim<sup>1</sup>, Jae Y. Lee<sup>2</sup>, Jae Doug Yoo<sup>3</sup>, Kwan H. Lee<sup>4</sup>  
GIST<sup>1 2 3 4</sup>

### 요약

본 연구의 목적은 객체의 형상정보(3D mesh)와 색/질감정보(image)를 이용하여 텍스처 맵핑된 고품질의 가상모델을 생성하는데 있다. 3 차원 형상정보에 대응하는 이미지 상의 텍스처 좌표 관계를 구하기 위해 오브젝트 좌표계와 카메라 좌표계 사이의 변환행렬, 카메라의 초점거리, 카메라 CCD 와 프레임상의 이미지 사이의 aspect ratio 를 파라미터로 하는 3D-2D 정합을 수행한다. 이러한 3D-2D 정합을 효율적으로 수행하기 위하여, 카메라 내부파라미터 검정단계, 신뢰도가 높은 초기해 설정단계, 비선형 최적화(Newton method) 단계로 접근한다. 또한, 색/질감정보로 이용되는 객체의 이미지는 촬영조건에 의해 스펙클러(specular)나 이미지 픽셀값의 포화상태(saturation) 등의 결점을 포함한다. 영상내의 스펙클러 좌표와 3D-2D 정합의 결과를 이용하여 촬영 당시의 광원을 추정하고, 근사화된 빛반사도 함수(BRDF)를 이용하여, 텍스처의 픽셀값 변조를 통해 이미지 촬영 당시의 광원효과가 제거된 디퓨즈 텍스처를 획득한다. 본 연구에서는 Phong의 모델을 근사화한 빛 반사도 함수 모델로 사용하였다.

Keyword : diffuse texture acquisition of a real object , camera calibration, HDRI, BRDF, 3D-2D registration

### 1. 서론

컴퓨팅 환경의 개선, 네트워크, 인터넷 기술의 발전 및 관련 인프라(infrastructure)의 구축으로 전자상거래는(e-Business)는 하나의 산업 분야로 자리매김 하고 있다. 산업구조의 커다란 변화를 가져온 전자상거래 산업으로 광고, 유통, 마케팅, 구매 등이 온라인 상에서 활성화 되어감에 따라 제품의 3 차원 형상 정보뿐만 아니라 색/질감 정보의 정확성이 요구 된다. 본 연구는 실세계 오브젝트의 3 차원 형상과 2 차원 이미지의 정합 그리고 이미지 픽셀값 변조를 통하여 고품질의 실감모델 생성을 목적으로 한다.

3 차원 형상에 대한 색/질감 정보를 획득하는 방법으로는 3 차원 컬러스캐너를 이용하여 형상과 색/질감 정보를 동시에 측정하는 방법과 이미지를

3 차원 형상정보에 매핑하는 텍스처 매핑 방법으로 구분 할 수 있다. 비전기반의 3 차원 컬러 스캐너의 경우, 고품질 대용량의 데이터를 효율적으로 획득할 수 있으나, 반사도가 높은 오브젝트는 측정이 힘든 단점이 있다. 본 연구에서는 카메라 내부파라미터 검정 과정, 3D-2D 정합의 신뢰도가 높은 초기해 설정과정, 비선형 최적화 과정을 통하여 효율적인 텍스처 획득 기법(texture acquisition)을 제안한다. 또한, 다양한 셔터스피드의 영상과 3D-2D 정합의 결과를 이용하여 촬영 당시의 광원을 추정하고, 근사화된 빛반사도 함수를 이용하여 촬영 당시의 광원효과가 제거된 디퓨즈 텍스처 획득 방법을 제안한다.

## 2. 관련 연구

실세계 오브젝트의 3D-2D 정합은 그 대상이 건물, 산업제품, 유물, 인간 등 다양한 분야에서 연구되어 왔다.

Guenter 는 마커를 이용하여 3D-2D 대응점을 자동으로 추출하는 방법을 제시하였으나, 텍스처정보에 마커가 포함되는 단점이 있다[1]. Debevec 은 대응점을 수작업을 통하여 지정해주는 가장 일반적인 방법을 사용하였고(주)아이너스 기술의 상용 역공학 모델러인 Rapidform 역시 이 방식을 사용한다[2]. 이 방식은 수작업의 특성상 정확도가 작업자에 의존적이다. David 등은 3D-2D 정합을 점이 아닌 3 차원 서피스(surface)정보를 이미지 평면에 투영시켜 두 개의 contour 사이의 오차를 최소화 하는 방법을 제시 하였으나, 정합의 정밀도가 이미지 내의 객체 추출 정밀도에 의존하는 단점이 있다[3]. Hendrik 은 2D 이미지와 비교할 3 차원 모델의 가상 이미지를 3D 모델의 가시화(rendering)을 통한 하드웨어 가속화 방식을 제안하였다[4].

## 3. 3D-2D 정합

3D-2D 정합은 오브젝트 좌표계와 카메라 좌표계사이의 변환 파라미터와 카메라 내부 파라미터의 최적화 문제로 설명 할 수 있다. 효율적인 최적화를 위하여 카메라 검정단계와 비선형최적화 단계로 나누어 카메라 내부 파라미터와 좌표계 변환 파라미터를 각각 구하는 방법으로 접근한다.

### 3.1 카메라 변환

본 연구에서는 그림 1 과 같이 Tsai 가 제안한 카메라 모델을 이용하였다[5]. 월드좌표계를  $(X_w, Y_w, Z_w)$ 로, 카메라 좌표계를  $(x, y, z)$ 로, CCD 좌표계를  $(u, v)$ , 컴퓨터 프레임 상의 이미지 좌표계를  $(U, V)$ 로 둘 때, 핀홀 카메라 모델에 의하여 이들 좌표계 사이의 관계는 각각 식(1), (2), (3)과 같다.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R_c \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T_c \quad (1)$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

$$X_f = s d'_x U_d + C_x, \quad Y_f = d'_y V_d + C_y, \quad (3)$$

$$d'_x = d_x N_{cx} / N_{fx}, \quad d'_y = d_y N_{cy} / N_{fy}$$

여기서,  $R_c$  는 3x3 회전행렬,  $T_c$  는 변이벡터,  $f$  는 초점 거리,  $s$  는 CCD 와 이미지 사이의 aspect ratio,  $(C_x, C_y)$ 는 CCD 중심점에 대응되는 이미지 좌표,  $d_x, d_y$ 는 각각 u 방향, v 방향 CCD cell 하나의 크기,  $N_{cx}, N_{cy}$ 는 각각 u 방향, v 방향 CCD cell 들의 크기,  $N_{fx}, N_{fy}$  는 각각 U 방향, V 방향 이미지 픽셀들의 개수 이다.

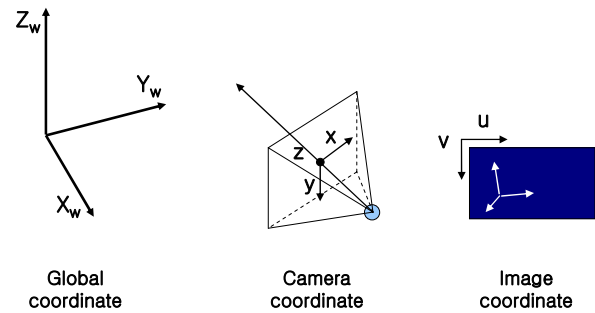


그림 1. Camera transformation

### 3.2 카메라 내부파라미터 검정

카메라 내부 파라미터( $f, s$ )는 그림 2 의 보정 패턴을 이용하여 Tsai 의 카메라 검정법으로 결정 하였다[5]. 보정패턴 특징점들의 월드좌표계 좌표값과 이미지좌표계 좌표값을 식(1), (2), (3)에 대입하여 생성된 행렬을 SVD(singular value decomposition) 접근법으로 푼다[6].

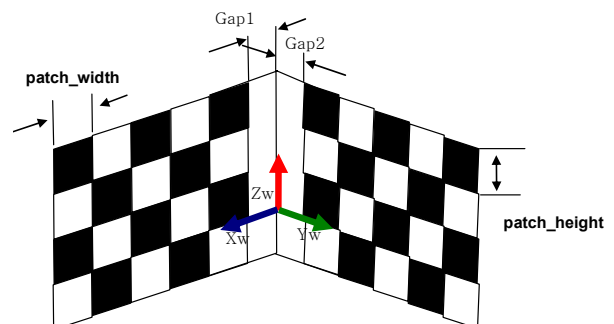


그림 2. Calibration pattern

### 3.3 좌표계 변환 파라미터 검정

오브젝트 좌표계에서의 객체가 카메라의 좌표계에서 영상내의 객체와 일치되는 회전변환 행렬을 R 이라 할 때, x, y, z 축에 대한 회전각  $\alpha, \beta, \gamma$ 를 이용하여 식(4)와 같이 회전변환 행렬 R 을 일반화 할 수 있다.

$$R = R_x^\alpha R_y^\beta R_z^\gamma \quad (4)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\alpha & s\alpha \\ 0 & -s\alpha & c\alpha \end{bmatrix} \begin{bmatrix} c\beta & 0 & -s\beta \\ 0 & 1 & 0 \\ s\beta & 0 & c\beta \end{bmatrix} \begin{bmatrix} c\gamma & s\gamma & 0 \\ -s\gamma & c\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} c\beta c\gamma & c\beta s\gamma & -s\beta \\ s\alpha s\beta c\gamma - c\alpha s\gamma & s\alpha s\beta s\gamma + c\alpha c\gamma & s\alpha c\beta \\ c\alpha s\beta c\gamma - s\alpha s\gamma & c\alpha s\beta s\gamma - s\alpha c\gamma & c\alpha c\beta \end{bmatrix}$$

그림 3 과 같이 임의의 회전 행렬 R 이 선택되면, 영상내 객체의 바운딩 박스를 형성하는 최대, 최저의 특징벡터  $\vec{t}_1, \vec{t}_2$  를 이용하여 3D 객체의 가상 투영 바운딩 박스를 구할 수 있다.

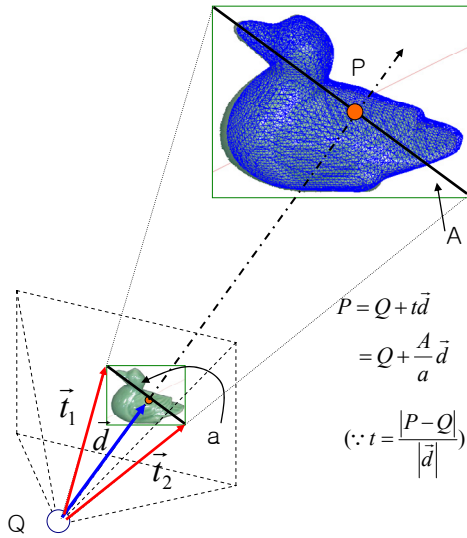


그림 3. Condition of 'Translation Vector decision'

3D 객체의 가상 투영 바운딩 박스 중심점과 가상 영상 객체의 중심점이 일치 해야 하는 구속조건과 이 두 영상의 오브젝트를 감싸는 바운딩 박스(bounding box)의 크기가 일치 해야 하는 구속조건으로 변이 벡터 T 는 자동으로 결정된다. 즉, 최적의 자세를 구하는 문제와 일치하게 된다. (이 두가지 구속조건과 식(1)-(4)으로부터 가상투영 영상을 얻을 수 있고, 회전

행렬 R 이 최적해 일 때, 실제영상과 가상 투영영상은 최소오차를 가지게 된다.)

비선형 최적화 문제를 풀기 위하여 식(5)와 같이 비용함수(cost function)를 설정한다.

$$\text{Minimize } J = \sum_i \sum_j e_{i,j} \quad (5)$$

$$\text{where } e_{i,j} = \text{XOR}(I_{i,j}, I'_{i,j})$$

여기서, XOR 은 'Exclusive OR' operation, i, j 는 이미지 가로세로 인덱스, I 는 촬영된 영상을 배경과 오브젝트에 대해 바이너리로 표현한 영상, I' 은 가상 투영 바이너리 영상이다. 본 연구에서 영상의 배경과 오브젝트의 분리는 블루스크린 기법(blue screen)을 이용하였다. 마우스를 이용하는 유저 인터랙션(user interaction)과 식(4)의 inverse kinematics 접근법으로 최적해에 가깝게 초기해를 설정하고,  $\partial J / \partial \alpha, \partial J / \partial \beta, \partial J / \partial \gamma$  로부터  $\Delta \alpha, \Delta \beta, \Delta \gamma$ 를 결정하는 뉴턴방법(Newton method)을 이용하여 비선형 최적해를 구하였다[6].

## 4. 디퓨즈 텍스처

카메라로부터 획득되는 영상은 촬영조건과 8bit 양자화(quantizing)에 의해 스펙클러(specular)나 픽셀값의 포화상태(saturation) 등의 현상을 포함한다. 가상 객체의 텍스처에 이러한 현상이 포함되어 있으면, 시각화(rendering)시 부자연스러움을 주는 결점의 요인이 된다. 영상의 촬영 당시의 광원효과를 제거하기 위하여, 영상내의 스펙클러 좌표와 3D-2D 정합 결과를 이용하여 영상 촬영시의 광원을 추정하고, 추정된 소수광원방향으로부터 픽셀값을 변조한다.

### 4.1 방향성 광원 추정

본 연구에서 이미지 상의 R, G, B 값이 (250, 250, 250) 이상이면, 빛이 정반사에 의한 스펙클러 현상으로 가정하였다.

그림 4 와 같이 가정에 의해 스펙클러로 선택된 픽셀의 이미지 좌표와 3D-2D 정합의 결과를 이용하여 영상의 광원방향을 추정한다.

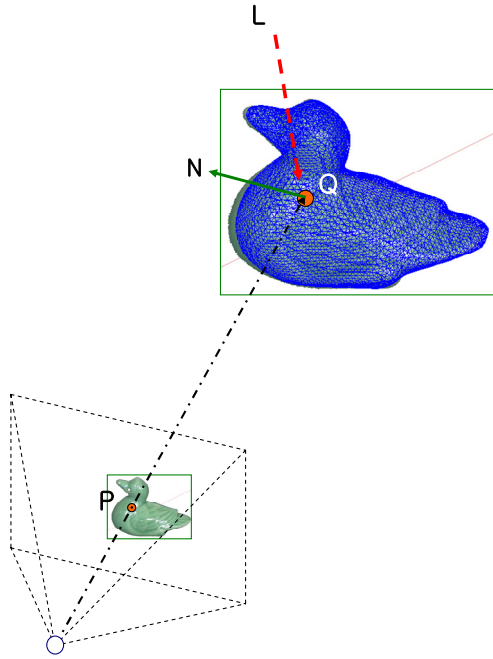


그림 4. 광원 방향성 추정

여기서 점 P 는 이미지상의 추정된 스페큘러 픽셀, 점 Q 는 3 차원 객체 상의 스페큘러 위치, 벡터 N 는 점 Q 에서의 3 차원 객체의 법선 벡터, 벡터 L 은 추정된 광원의 방향이다.

#### 4.2 픽셀값 변조

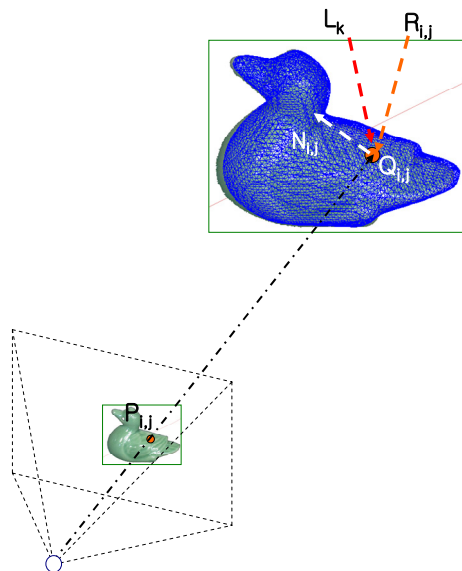


그림 5. 이미지 내 의 픽셀 정반사 벡터 추정

그림 5 과 같이, 추정된 광원 방향 단위벡터를  $L_k$  ( $k= 0, \dots, m$ ), 이미지 픽셀을  $P_{i,j}$ ,  $P_{i,j}$  에 대응하는 3 차원 객체의 표면위의 좌표를  $Q_{i,j}$ , 카메라의 위치 O 에서 CCD 상의 이미지 픽셀  $P_{i,j}$  를 잇는

벡터와 3 차원 객체의 좌표  $Q_{i,j}$  의 법선벡터에 대하여 정반사되는 방향 단위벡터를  $R_{i,j}$  이라 두었을 때, 픽셀값 변조를 위한 가중치  $F_{i,j}$  는 식 (6)에 의해 결정된다.

$$F_{i,j} = 1 + 0.2 * (t_{i,j} - A) / V \quad (6)$$

$$t_{i,j} = \sum_{k=0}^m (R_{i,j} \cdot L_k)^n$$

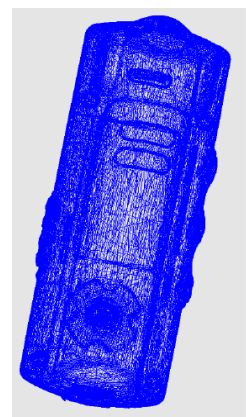
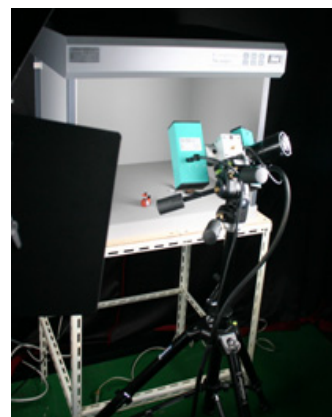
$$A = Avg(t_{i,j})$$

$$V = Var(t_{i,j})$$

여기서, n 은 phong's highlighting 의 shiness factor, Avg 는 mean operation, Var 는 variance operation 을 뜻한다. 본 연구에서는  $n=3$  으로 가정하였다.

#### 5. 실험 및 결과

독일 Breuckman 社의 3D 형상 측정기인 optoscan HE-100 를 이용하여 3 차원 점데이터를 획득하고, ㈜ 아이너스 기술社의 3D 역형상공학 모델러인 Rapidform2004 를 이용하여 모델링 하였다. 그림. 6 의 (a)는 3D 점데이터를 획득하는 스캐닝시스템을 (b)는 MP3 플레이어의 획득된 점데이터로부터 모델링한 3 각 메쉬(mesh)를 나타낸다.



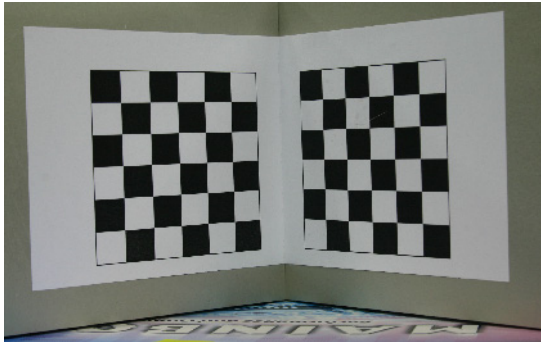
(a)형상 측정 시스템

(b) Mesh model

그림 6. 3D geometry data acquisition

카메라 검정 및 텍스처 정보를 위한 영상은 Canon 社의 EOS 300D 카메라를 이용하여 촬영하였다 그림 7 의 (a)는 카메라 내부

파라미터를 검정하기 위해 촬영한 패턴 영상이고, (b)는 색/질감정보를 얻기 위한 객체 이미지, (c)는 관심 오브젝트만 추출한 영상이다



(a) Calibration pattern



(b) Mp3 player image



(c) Segmented object

그림 7. Camera & Texture Data Acquisition

그림 8 은 2D-3D 정합과정을 보여준다.

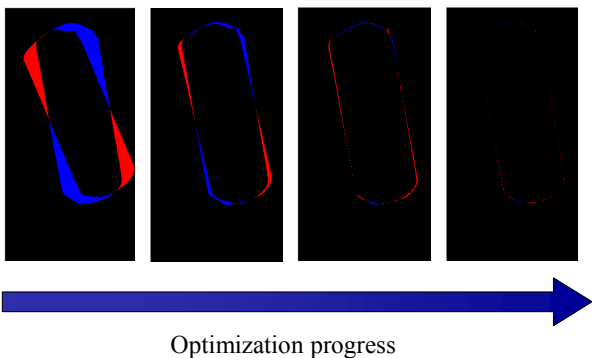


그림 8. Optimization process

그림 9 는 2D-3D 정합결과 획득된 텍스처 맵(a)과 픽셀변조 결과 획득된 디퓨즈 맵(b)을 보여준다.



(a) texture map



(b) diffuse map

그림 9. 픽셀 변조 결과

그림 10 은 삼각망 12 만개의 mesh model 과 전체크기 3072\*2048 (이미지 내의 객체 크기 :718\*1103) 의 3D-2D 정합의 렌더링 결과와 픽셀 변조를 통한 디퓨즈 텍스처를 이용한 렌더링 결과를 보여준다.

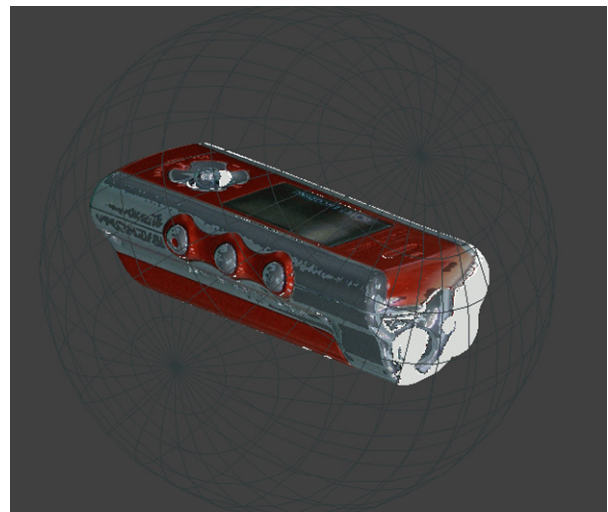
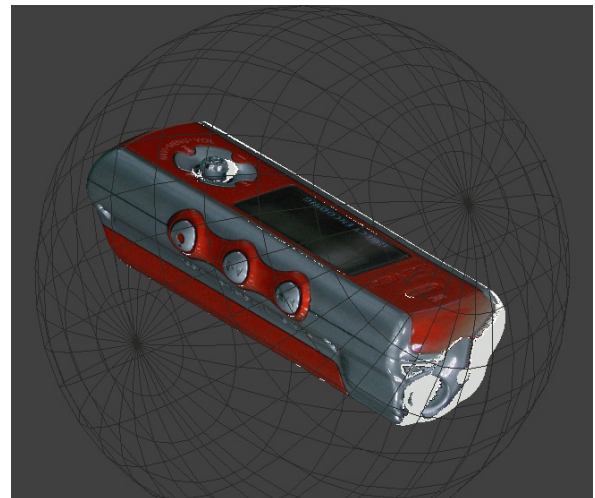


그림 10. Rendering result

## 6. 결론 및 향후 연구 과제

실제 존재하는 오브젝트의 형상에 대응하는 색/질감을 획득하는 효율적인 방법을 제시하였다. 이로써, 형상 획득과 텍스처 획득을 서로 독립적으로 측정 가능하게 되어 데이터 획득의 유연성을 높였다. 또한, 카메라 검정과 3D-2D 정합의 두 단계 접근법과 인터페이싱을 통한 적절한 초기해 설정으로, 비선형 최적화 문제를 좀 더 간편히 하였다. 또한, 광원의 추정과 이를 기반한 이미지 촬영시 광원 효과를 줄이는 방법을 제시 하였다. 그러나 소수광원 의 선택문제와 근사화된 빛반사도 문제 등으로 부드럽지 못한 결과를 얻게 되었다. 메쉬 하나당 여러 시점의 영상으로부터 획득되는 다수의 텍스처 정보의 융합, 변조 등의 연구가 향후 수행되어야 한다.

### 감사의 글

본 연구는 광주과학기술원 실감방송 연구센터 를 통한 정보통신부 대학 IT 연구센터(ITRC) 육성 사업의 지원과 광주과학기술원 실감콘텐츠 연구 센터(ICRC)를 통한 과학기술부 사업의 지원에 의한 것입니다.

### 참고 문헌

- [1] B. Guenter, C. Grimm, D. Wood, H. Malvar, F. Pighin, "Making Faces," *Proceedings of SIGGRAPH98*, pp55-66, July 1998.
- [2] P. E. Debevec, C. J. Taylor, and J. Malik, "Modeling and Rendering Architecture from Photographs : A Hybrid Geometry- and Image-based Approach," *Proceedings of Computer Vision (ECCV '92)*
- [3] David G. Lowe, "Fitting Parameterized Three-Dimensional Models to Images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 5, May 1991.
- [4] Hendrik P. A. Lensch, Wolfgang Heidrich, Hans-Peter Seidel, "Automated Texture Registration and Stitching for Real World Models," *Proceedings of the 8th Pacific Conference on Computer Graphics and Applications*, pp317-452, 2000.

[5] Roger Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4, August 1987.

[6] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, *Numerical Recipes in C++*, 2nd edition Cambridge university press