사진 속 피사체의 법선 벡터 예측에 기반한 햅틱 상호 작용

Haptic Interaction with Objects Displayed in a Picture based on Surface Normal Estimation

김 승 찬¹, 권 동 수²

Seung-Chan Kim¹, Dong-Soo Kwon²

Abstract In this paper we propose a haptic interaction system that physically represents the underlying geometry of objects displayed in a 2D picture, i.e., a digital image. To obtain the object's geometry displayed in the picture, we estimate the physical transformation between the object plane and the image plane based on homographic information. We then calculate the rotated surface normal vector of the object's face and place it on the corresponding part in the 2D image. The purpose of this setup is to create a force that can be rendered along with the image without distorting the visual information. We evaluated the proposed haptic rendering system using a set of pictures of objects with different orientations. The experimental results show that the participants reliably identified the geometric configuration by touching the object in the picture. We conclude this paper with a set of applications.

Keywords: Image Understanding, Resistive Force, Force Display

1. 서 론

사용자들은 사진을 통해 주변의 다양한 일상을 기록한다. 컴퓨팅 환경이 발전함에 따라 이러한 사진은 디지털화 된 정보로서 저장, 활용이 되고 있다. 최근 시각 정보에 사용자가 만질 수 있는 햅틱 정보를 추가하는 연구가 많이 이루어지고 있는데, 주로 외부 측정 장치 혹은 저작 도구를 이용한 방식이 활용되고 있다. 시/공간적인 정보의 해상도 문제를 해결하고자 파라미터로 햅틱 텍스처/기하 정보를 동적으로 생성하는 절차적(procedural) 방식이 적용

이 되기도 하지만^[1-3], 파라미터 튜닝과 앨리어싱 등의 이슈를 가지고 있어 그 활용 범위가 제약이 되고 있다.

가장 널리 활용되고 있는 이미지 기반의 음영정보-기하 정보 사이의 맵핑 방법은 진동장치^[4], 질감제시장치^[5], 힘제시 장치^[6,8] 등에 걸쳐 다양한 응용 분야에 활용이 되고 있다.

이와는 달리 시각-햅틱정보 사이의 미리 정해진 맵핑을 통해 주변의 환경에 촉각 정보를 할당하는 방법 또한 활용이 되고 있다. 최근의 한 연구에서는 특정 색에 미리 정의된 햅틱 정보를 맵핑하여 사용자에게 주변에 관련한 유용한 정보를 전달할 수 있음을 실험을 통해 검증한 바 있다^[9].

본 논문에서는 시각 정보에 촉각 정보를 추가하려는 기존 연구의 연장선 상에서 사진 속 관심 물체의 macro-level 기하 정보에 초점을 두고, 이를 예측, 햅틱 렌더링하는 시스템을 제안하고 검증한다. 사진 속 피사체를 평면 표면 혹은 평면 표면으로 분해될 수 있는 물체 (예, 건물 의 벽, 곡면)로 한정하고 연구를 진행한다.

Received : Jun. 3. 2013; Reviewed : Jun. 25. 2013; Accepted : Jul. 29. 2013

※ 본 연구는 지식경제부 및 한국산업기술평가관리원의 IT 신성장 동력핵심기술개발 사업의 일환으로 수행하였음. [2009-F-048-01, 비접촉식 멀티포인트 질감 인터랙션 기술 개발]

※ This work was supported by the IT R&D program of MKE/KEIT, [2009-S-035-01, Contact-free Multipoint Realistic Interaction Technology Development].

¹ Mechanical Engineering, KAIST (kimsc@robot.kaist.ac.kr)

² Mechanical Engineering, KAIST (kwonds@kaist.ac.kr)

† Corresponding author: Mechanical Engineering, KAIST, Daehak-ro, Yuseong-gu, Daejeon, Korea (kwonds@kaist.ac.kr)

2. 일반적인 2D 이미지와의 햅틱 상호작용

2.1. 개요

본 챗터에서는 2D 시각정보에 적합한 햅틱 정보를 정의하고 이를 카메라 보정 기법을 활용하여 추출하는 방법론을 다룬다. 2D 시각 정보에 적합한 햅틱 정보라 함은 사진으로부터 3차원 시각 정보를 재생성하는 방법론에 기반하는 것이 아닌, 2D 이미지의 시각적 특성을 보존하면서 이에 자연스럽게 중첩될 수 있는 촉각 정보를 의미한다. 제안하는 방법은 사진 속 물체의 기하학적 배치 - 물체의 평면과 이미지 평면 사이의 상대적인 회전 정보 - 를 예측하는 것으로부터 시작한다. 카메라 보정 과정에서 이 물리적 변환 (physical transformation)은 외부파라미터 (extrinsic part)로 정의됨에 따라 이하 두 용어는 상황에 따라 번갈아가며 사용될 수 있음을 미리 알려둔다.

이렇게 얻어진 기하학적 회전 정보는 사용자가 사진의 특정 부분을 만질 때 저항력(resistive force)을 발생시키는 용도로 사용이 된다. 이러한 저항력 기반의 햅틱 렌더링 방법은 1990년대 Minsky에 의해 처음 소개된 바 있다^[8]. 이 방법은 햅틱 상호작용 포인터(HIP)가 위치하고 있는 지점의 기하학적 gradient 정보에 기반하여 HIP이 낮은 위치로 장치가 위치하도록 횡방향 힘 (lateral force)을 전달하는 방식으로 요약해 설명할 수 있다. 이를 통해 z -방향의 힘을 사용하지 않고, xy 평면에서의 힘만 방향으로 사용자에게 z -방향의 기하 정보가 존재하는 것 같은 착각 현상을 일으키도록 하였다. 이러한 방법론은 이후 다양한 응용분야에 적용된바 있다.^[10-12]

같은 맥락에서 본 논문에서는 디지털 이미지의 2D 속성을 보존하며 햅틱 피드백을 전달하기 위해 횡방향 힘을

도입/활용하였다. 아래 그림은 사용자가 2D 이미지 평면에 디스플레이 되고 있는 물체를 만질 때, 물체 평면의 보여지는 기울기에 근거하여 힘 정보를 생성하는 개념도를 보여주고 있다.

2.2 두 평면의 물리적 변환 관계 예측

두 평면 사이의 물리적 변환은 원 평면의 위치 예측 과정(pose estimation process)에 사용되거나 HCI 관점에서 다양한 응용 분야에 활용이 되기도 한다^[13,14].

물체의 좌표를 $\mathbf{X} (\in \mathbb{R}^3)$ 라 정의하고, 해당 점의 이미지 공간에 해당하는 점을 $\mathbf{x} (\in \mathbb{R}^2)$ 라 할때, 각각을 homogeneous 좌표계에서 $\tilde{\mathbf{X}} (\in \mathbb{P}^3)$, $\tilde{\mathbf{x}} (\in \mathbb{P}^2)$ 으로 표현한다면 두 평면 사이의 물리적 관계식은 다음과 같이 정리될 수 있다.

$$\tilde{\mathbf{x}} = sH\tilde{\mathbf{X}} = sK[R|\mathbf{t}]\tilde{\mathbf{X}} \quad (1)$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = s \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

이때 s 는 스케일 요소, K 는 카메라 행렬 (camera matrix)이며, 초점거리(focal lengths)인 f_x 와 f_y 그리고 이미지의 중심점에 해당하는 c_x 와 c_y 로 구성되어 있다.

물리적 변환 행렬은 3×3의 회전 행렬 R 과 3×1의 변위 벡터 $\mathbf{t}$ 로 구성되어 있다. 즉 호모그래피 행렬(homography matrix), H 는 카메라 행렬, K 과 물리적 변환 행렬 $[R|\mathbf{t}]$ 로 구성이 되며, 위와 같은 관계식으로 물체의 좌표 $\tilde{\mathbf{X}}$ 를 $\tilde{\mathbf{x}}$ 로 변환한다.

2.3 법선 벡터 필드 생성

주어진 평면의 물리적 변환 행렬의 회전 정보 R 을 통해 다음과 같이 이미지 공간에 법선 벡터 필드를 생성한다.

$$\tilde{\mathbf{n}}_T = W\tilde{\mathbf{n}}$$

$$W = \begin{bmatrix} R & \mathbf{t} \\ 0_3^T & 1 \end{bmatrix} = \begin{bmatrix} R & 0_3 \\ 0_3^T & 1 \end{bmatrix} \quad (2)$$

단, $\tilde{\mathbf{n}}$ 는 4×1의 본래 법선 벡터, $\tilde{\mathbf{n}}_T$ 는 회전된 4×1의 법선벡터임.

Fig. 1. Haptic rendering method based on the viewed slope

위의 결과 생성된 회전된 법선 벡터는 2D 이미지 공간에 배치되어 2.1 절에서 설명된 바와 같이 x, y 축의 반력을 생성하는데 사용이 된다.

2.4 햅틱 렌더링

본 챕터에서는 예측된 기울어진 형태로 사진 속에 보이는 평면의 법선 벡터를 2.1절에서 설명된 바와 같이 횡방향 힘을 통해 전달하는 방법론을 설명한다. 즉, 미리 계산된 법선 벡터 필드 위에서 사용자가 움직일 때의 상호 작용을 정량화한다. 이를 위해 우선 사용자의 움직임을 다음과 같이 단위 방향 벡터로서 정의한다.

법선 벡터로 특성화된 사진 속 평면에서 사용자의 움직임은 다음과 같이 두 벡터 사이의 내적으로 정의한다. 이는 프로젝션의 개념으로 사용자가 물체의 법선 벡터 방향으로 움직인 정도를 횡방향 힘움직였는지의 정도를 가리키는 물리적 의미를 내포하고 있다.

$$F(x, y) = \mathbf{n}_T \cdot \hat{\mathbf{m}} \quad (3)$$

위 연산의 부호에 따라 힘의 방향과 크기가 결정이 된다.

Fig. 2. Unit direction vector for defining the user motion

Fig. 3. Physical sense of the slope-based haptic interaction

3. 평가

3.1 실험 자극

실험을 위해 한 변이 20cm인 정육면체 상자가 사진 속

물체 (피사체)로서 활용이 된다. 다음 그림에서 보이는 바와 같이 물체를 5도 간격으로 15도에서 75도 사이의 값으로 y 축 회전이 된 상태로 사진 촬영을 하였다.

실험을 위해 각 면의 꼭지점, $\mathbf{x} (\in \mathbb{R}^2)$ 은 이미지 프로세싱에 의존하지 않고 실험자에 의해 수동으로 선택 (picking)이 되었다. 이는 실험을 위한 예측 오차를 줄이기 위함인데, 예측된 법선 벡터와 계산에 의한 값과의 비율은 다음과 같다 : 0.99 (x 축 왼쪽), 1.03 (z 축 왼쪽), 1.02 (x 축 오른쪽), 0.99 (z 축 오른쪽).

Fig. 4. A set of pictures varying in terms of their orientation on the y -axis

3.2 실험 장치

3.2.1 힘 제시 장치

마우스 커서 입력 장치 및 정량적 힘 제시 장치로서 SensAble PHANToM Omni TM를 사용하였다. 2차원 커서 제어를 위해 6.4 px/mm의 비율로서 입력 장치가 사용이 되었다.

3.2.2 응답 장치

피험자의 답변을 정량화 하기 위해 화면에 보이는 사진과 동일한 형태의 탠저블 인터페이스를 개발하였다. 해당 장치의 손잡이 부분은 정육면체 형태를 취하고 있으며, 사진과 동일한 축으로 회전이 가능하도록 설계되었다. 즉 사용자는 실험과정 중 느낀 정도를 개발한 인터페이스를 회전 시킴으로써 응답을 하게 되고 장치 (MX-28, Robotis Co.)의 엔코더에서 이 값을 측정함으로써 제출된 응답을 기록하게 된다.

Fig. 5. Developed tangible interface

3.2.3 실험 단계

실험은 두개의 다른 그룹으로 나뉘어 피험자간 설계 (between-subject design) 형태로 실험이 진행되었다.

시각 자극 그룹 (V)

본 그룹은 햅틱 정보는 받지 않고 화면 속에 표시된 사진을 보며 그림 5의 인터페이스를 통해 답변을 한다. 사람이 시각을 통해 물체의 배치를 쉽게 파악한다는 점에 근거하여^[15], 다음의 H그룹의 비교군으로 설정하였다. 각각의 자극은 6번 반복되어 전체 실험은 78 (=13 × 6) 개의 trial로 구성이 된다.

햅틱 자극 그룹 (H)

V 그룹과 사용된 자극은 동일하며, 자극이 시각정보로 전달되지 않고, 힘 제시 장치를 통해 전달된다. 사용된 힘 정보는 수식 3에 근거하여 ± 0.6 N 사이의 값으로 결정되었다.

두 그룹 모두 본 실험 후 NASA TLX 질문지를 통해 실험시의 work load를 측정하여 시스템의 적합성을 측정하였다.

3.2.4 피험자

본 실험의 V그룹과 H그룹에참여한 피험자는 각각 12명 (남자 7명), 14명(남자 8명)이며 평균 나이는 각각 21.9세, 21.7세의 대학교 학생들로 구성이 된다. 전원 오른손 잡이로서, 실험에 결격이 되는 알려진 cutaneous 또는 kinesthetic 문제가 없는 집단으로 구성되었다.

3.2.5 실험 결과

전체적인 에러의 절대값은 V그룹의 경우 평균 5.5도 (표준편차 3.8도)이며, H그룹의 경우 8.2도(표준편차 6.5도)이다. 또한 각각의 trial 수행 시간은 V그룹의 경우 5.06초, H그룹의 경우 12.59초로 측정되었다.

일원분산분석 (One-way ANOVA test) 결과 두 그룹 사이의 에러 절대값에 유의미한 차이가 있음이 확인되었다. 다음의 그림은 각 그룹별 각도별 응답 결과를 보인다.

H그룹의 에러가 V그룹에 비해 1.5배 크게 측정이 되었지만, 시각 정보가 물체의 배치 정보를 보다 효율적으로 측정한다는 점을 고려해 본다면^[15], 피험자들이 햅틱 정보에만 의존하고도 사진 속 물체의 회전 정보를 전반적으로 용이하게 파악하고 있다고 해석해 볼 수 있다. TLX응답 결과 V와 H그룹 각각 100점 기준으로 22.9 와 38.3 로 측정되었는데, 실험 응답 결과와 종합하여 볼 때 제안하는 방법론이 실제적으로 적용가능함을 확인할 수 있다.

Fig. 6. Responses with respect to the experimental condition

4. 응용 분야

본 챕터에서는 예측된 기하 정보에 기반한 2D 이미지와의 햅틱 상호작용에 대해 논의한다.

4.1 모서리 렌더링

물체의 모서리는 일상에서 물체를 인식하는데 중요한 역할을 하는 특징점 중 하나이다^[16]. 따라서 많은 햅틱 인터페이스 시스템이 이미지 속 물체의 모서리를 다양한 맥락에서 렌더링한다. 그러나 3차원 기하학 정보가 주어지지 않는 경우 이미지 프로세싱에 의존하는 경우가 많은

Fig. 7. Example of a 2D picture with estimated surface normals displayed

데, 이 경우 기하 정보의 변화에 의해 생긴 모서리 뿐 아니라 다른 요소들 (예, 표면 색상 불연속성, 그림자) 에 의해 생성된 불연속 점까지 모서리의 범주에 들게되는 단점이 존재한다. 또한 햅틱 정보가 모서리에서만 전달된다는 점은 실제 일상 생활에서 얻는 물체의 느낌과는 다른 큰 차이를 만들어내게 된다.

본 논문에서는 이러한 단점을 각각의 평면에서 횡방향 힘을 전달해주어 저항감을 전달하는 방식으로 해결한다. 즉, 인접한 표면으로 커서가 이동시 불연속적인 법선벡터가 만들어내는 힘정보의 불연속성으로 모서리 느낌을 재현해 낼 수 있게 되는 것이다. 다음의 그림은 제안하는 햅틱 렌더링 방법을 위해 생성된 표면 법선 벡터 필드를 보이고 있다.

4.2 Scattering and Feeling

본 논문에서 제안하는 방법론이 실질적으로 활용이 되기 위해서는 이미지 프로세싱에 의해 결정되는 $\mathbf{x} \in \mathbb{R}^2$ 와 사용자가 시스템에 입력으로 넣어주는 모델 정보 (예, 평면의 길이 정보)인 $\mathbf{X} \in \mathbb{R}^3$ 이 필요하게 된다. IR마커 등을 이용해 실질적으로 해당 정보를 용이하는 증강현실의 방법이 대안으로 사용될 수 있다.

본 챕터에서는 해당 햅틱 상호작용 방법을 사용자 주변을 만져볼 수 있도록 하는 기존연구 [17] 에서 착안하여, 미리 정해진 평면의 정보 ($\mathbf{X} \in \mathbb{R}^3$) 를 투사하고 카메라로 이를 읽어 이미지 프로세싱을 통해 $\mathbf{x} \in \mathbb{R}^2$ 를 결정하여 주변 환경의 법선 벡터를 예측하는 방법론을 제안한다. 아래 그림은 일정간격 5×4의 grid 점들을 모바일 프로젝터

Fig. 8. Physical interaction with objects using a scattered array of points

로 주변 공간에 뿌리고 카메라를 통해 되어어진 영상 정보를 통해 주변 공간(곡면)을 유한개의 평면의 조합으로 근사하여 렌더링하는 예를 보이고 있다.

특히, 주변 공간이 곡면인 경우 force shading에 활용되는 법선 벡터 보간법^[18,19]을 2D 이미지 공간에 도입하여 투사하는 점의 숫자를 줄일 수 있게 된다. 아래는 해당 보간법을 이용하여 그림 8의 표면 (중央이 들어간 오목한 곡면) 왼쪽에서 오른쪽으로 만지며 지나갈 때의 힘 정보를 보이고 있다. 벡터 보간법 도입을 통해 인접한 면의 불연속적인 모서리 효과를 제거한다.

Fig. 9. Example of interpolated forces while moving across the surface (left to right)

5. 논의

수식 (1)로부터 예측된 변위 벡터 $\mathbf{t}$ 는 수식 (2)에서 힘을 생성하는데 활용이 되지 않는다. 이는 거리 정보, $\mathbf{t}$ 가 원

근 변환 (perspective transformation)에 의해 피사체의 크기에 이미 반영이 되었기 때문이다. 거리를 힘 정보로 맵핑하는 연구는 본 연구에서는 다루지 않는다.

6. 결 론

본 논문에서는 사진 속 피사체와의 상호작용을 위해 유효한 기하 정보를 법선 벡터 형태로 예측하여 추출하고, 이를 횡 방향 힘으로 사용자에게 제시한다. 횡방향 힘을 사용한 이유는 햅틱 렌더링 시 2D 이미지 특성을 보존하면서 이미지 표면에서 힘 정보를 제시하기 위함이다. 제안하는 방법론을 평가하기 위해 본 논문에서는 사용자가 사진 속 물체를 만지고 만진 물체의 회전 정보를 답하는 사용자 테스트를 수행하였다. 실험 결과 피험자들은 횡방향 힘에 근거하여 평면 물체의 회전 정보를 예측해 내었는데, 이를 통해 제안된 방법론이 2D 이미지에 햅틱 특성을 추가하는데 실질적으로 쓰일 수 있음을 확인하였다. 향후 연구로서, 제안하는 시스템을 소형화하여 사용자 주변의 환경을 모바일 장치를 통해 만질 수 있도록 하는 시스템을 개발할 예정이다.

References

- [1] J. P. Fritz, and K. E. Barner, "Stochastic models for haptic texture." pp. 34-44.
- [2] J. Shopf, and M. Olano, "Procedural haptic texture." pp. 179-185.
- [3] S.-C. Kim, S.-C. Kang, and D.-S. Kwon, "Sound Generation for the Haptic Perception using an Irregular Primitive Function." pp. 19-24.
- [4] P. Bach-y-Rita, C. C. Collins, F. A. Saunders *et al.*, "Vision substitution by tactile image projection," *Nature*, vol. 221, no. 6845, pp. 963-964, 1969.
- [5] M. Shinohara, Y. Shimizu, and A. Mochizuki, "Three-dimensional tactile display for the blind," *Rehabilitation Engineering, IEEE Transactions on*, vol. 6, no. 3, pp. 249-256, 1998.
- [6] B.-K. Han, S.-C. Kim, S.-C. Lim *et al.*, "Physical mobile interaction with kinesthetic feedback." pp. 571-575.
- [7] S.-C. Kim, B.-K. Han, and D.-S. Kwon, "Feeling the unseen: physical interaction with depth-embedded images," in CHI '13 Extended Abstracts on Human Factors in Computing Systems, Paris, France, 2013, pp. 727-732.
- [8] M. Minsky, O.-y. Ming, O. Steele *et al.*, "Feeling and seeing: issues in force display," *SIGGRAPH Comput. Graph.*, vol. 24, no. 2, pp. 235-241, 1990.
- [9] A. Israr, O. Bau, S.-C. Kim *et al.*, "Tactile feedback on flat surfaces for the visually impaired," in Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts, Austin, Texas, USA, 2012, pp. 1571-1576.
- [10] G. Robles-De-La-Torre, and V. Hayward, "Virtual surfaces and haptic shape perception." pp. 1081-1085.
- [11] G. Robles-De-La-Torre, and V. Hayward, "Force can overcome object geometry in the perception of shape through active touch," *Nature*, vol. 412, no. 6845, pp. 445-448, 2001.
- [12] S. Saga, and K. Deguchi, "Lateral-Force-Based 2.5-Dimensional Tactile Display for Touch Screen." pp. 15 - 22.
- [13] K. Hinckley, M. Sinclair, E. Hanson *et al.*, "The videomouse: a camera-based multi-degree-of-freedom input device." pp. 103-112.
- [14] K. D. D. Willis, I. Poupyrev, S. E. Hudson *et al.*, "SideBySide: ad-hoc multi-user interaction with handheld projectors," in Proceedings of the 24th annual ACM symposium on User interface software and technology, Santa Barbara, California, USA, 2011, pp. 431-440.
- [15] S. J. Lederman, and R. L. Klatzky, "Haptic identification of common objects: Effects of constraining the manual exploration process," *Attention, Perception, & Psychophysics*, vol. 66, no. 4, pp. 618-628, 2004.
- [16] S. J. Lederman, and R. L. Klatzky, "Haptic perception: A tutorial," *Attention, Perception, & Psychophysics*, vol. 71, no. 7, pp. 1439-1459, 2009.
- [17] H. Yano, Y. Miyamoto, and H. Iwata, "Haptic interface for perceiving remote object using a laser range finder," in Proceedings of the Third Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2009, pp. 196-201.
- [18] D. C. Ruspini, K. Kolarov, and O. Khatib, "The haptic display of complex graphical environments." pp. 345-352.
- [19] C. Basdogan, C. H. Ho, and M. Srinivasan, "A ray-based haptic rendering technique for displaying shape and texture of 3D objects in virtual environments." pp. 77-84.

김 승 찬

2005 연세대학교 기계공학
(공학사)

2007 한국과학기술원 기계공학
(공학 석사)

2013 한국과학기술원 기계공학
(공학 박사)

2011 미국 Disney Research Pittsburgh 인턴 연구원
관심분야 : Haptics, Robotics, HCI

권 동 수

1980 서울대학교 기계공학
(공학사)

1982 한국과학기술원 기계공학
(공학석사)

1991 미국 Georgia Institute of
Technology, Robotics
(공학박사)

1991~1995 미국 Oak Ridge 국립 연구소 선임 연구원
1995~현재 한국과학기술원 기계공학과 교수
관심분야 : HRI, Haptics, Medical Robots