

얼굴 방향에 기반을 둔 컴퓨터 화면 응시점 추적

오 승 환, 이 희 영
전남대학교 전자공학과

TEL: 062-530-0132, FAX: 062-530-1759

A Gaze Tracking based on the Head Pose in Computer Monitor

Seung Hwan Oh, Heyoung Lee

Dept. of Electronic Engineering, Chonnam National University

E-mail : neuron21@moiza.chonnam.ac.kr

Abstract

In this paper we concentrate on overall direction of the gaze based on a head pose for human computer interaction. To decide a gaze direction of user in a image, it is important to pick up facial feature exactly. For this, we binarize the input image and search two eyes and the mouth through the similarity of each block (aspect ratio, size, and average gray value) and geometric information of face at the binarized image. We create a imaginary plane on the line made by features of the real face and the pin hole of the camera to decide the head orientation. We call it the virtual facial plane. The position of a virtual facial plane is estimated through projected facial feature on the image plane. We find a gaze direction using the surface normal vector of the virtual facial plane.

This study using popular PC camera will contribute practical usage of gaze tracking technology.

I. 서론

최근 컴퓨터 기술의 향상과 다양한 멀티미디어 장치들의 발달과 함께 컴퓨터를 편리하게 사용하려는 욕구 또한 높아져 휴먼-컴퓨터 인터페이스에 대한 관심이 크게 부각되고 있다. 키보드나 마우스와 같은 기존의 입력 장치는 사람의 감정이나 의도와 같은 감성 정보를 컴퓨터에 입력시키기 어렵고, 사용자의 직접적인

입력이 없으면 컴퓨터 스스로 어떤 일을 하거나 판단을 내릴 수 없다.

이에 따라 기존의 수동적인 입력 장치를 대신하는 제스처나 얼굴 표정, 음성, 시선 등을 이용한 멀티 모달(Multi-Modal) 사용자 인터페이스에 대한 연구가 진행 중에 있다[9]. 그 중 마우스의 기능을 대신하여 사람의 응시 방향을 활용하는 방식은 컴퓨터가 사람의 의도를 인지한다는 측면에서 인간과 컴퓨터간의 새로운 의사소통 방식의 하나로 간주되며[8], 몸이 불편한 노약자 및 장애인들에게 편리하게 컴퓨터를 사용할 수 있도록 돕는 인간 친화 복지 기술이라 할 수 있다.

응시 위치를 결정하는 요소에는 머리의 움직임과 눈동자의 움직임이 있다. 머리의 움직임은 응시점의 전반적인 방향을 결정하고 눈의 움직임은 응시점의 정확한 위치를 결정한다[1]. 본 연구에서는 얼굴의 움직임을 이용한 대략적인 응시점 추적에 중점을 두었다.

본 논문에서는 얼굴의 응시 방향을 추정하기 위해 실제 얼굴을 대신하는 가상의 얼굴 평면(Virtual Facial Plane)을 설정하였다. 입력 영상에서 찾아낸 얼굴의 특징(두 눈과 입의 한쪽 끝점)을 통해 가상의 얼굴 평면을 이루는 세 점의 위치를 알아냄으로써 응시방향을 알아냈다. 가상의 얼굴 평면은 하나의 카메라를 가지고 사용자의 얼굴 크기나 위치, 얼굴 평면상의 세 점의 거리와 같은 기하학 정보(Geometric Information)를 정확히 알아내기가 어렵기 때문에 고정된 크기의 얼굴 평면을 설정하기 위한 것이다.

II. 얼굴 추적(Face Detection)

얼굴의 특징 점들을 찾고 추적하는 기술은 템플릿 매

칭(Template Matching)과 특징기반 방법(Feature based Method)으로 나눌 수 있다. 템플릿 매칭은 초기에 얼굴의 특징 점들에 해당하는 템플릿(Template)이 주어진 상태에서 그 템플릿과 가장 잘 맞는 영역을 선택하는 기술로써 초기에 적당한 템플릿이 주어지지 않으므로 사람마다 다른 템플릿을 설정해야 하는 (User-dependent) 단점이 있다[2][3]. 특징 기반 방법은 얼굴의 구성요소 (눈, 코, 입 등)들의 각각의 특징들과 서로간의 기하학 정보들(Geometric Information)을 이용하여 구하는 방법이다[4].

본 논문에서는 특징 기반 방법에 근거하여 눈과 입을 얼굴의 특징 파라메타로 정하고 Threshold값을 이용하여 입력 영상을 이진화 시킨 후, 이진화된 입력 영상에서 나타나는 어두운 점의 블록들 중에서 눈과 입을 찾아냈다. 특히, 두 눈에 해당하는 두 블록을 찾기 위해 각각의 블록의 유사도를 비교하는 완전그래프 매칭(Complete-Graph Matching)을 이용했다[1][6].

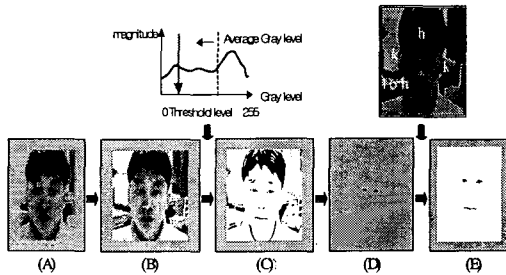


그림 3. Process to Find the Facial Feature

얼굴의 특징 점들을 찾아가는 순서는 그림 1.처럼 칼라 입력 영상(A)을 Gray level 이미지로 바꾸고(B) Threshold 값에 의해 Gray 이미지를 이진화 시킨다(C). 이진화된 영상에서 나타나는 블록들 중 각 블록의 종횡비, 평균 명암비, 크기 등의 유사도가 가장 큰 두 블록을 눈으로 하여 찾는다(D). 눈을 찾은 후, 두 눈 밑으로, 눈과 눈 사이에 입이 존재한다는 위치 정보를 이용하여 입에 해당하는 블록을 찾는다(E)[5].



그림 2. Detecting Facial Feature

III. 응시추적(Gaze Tracking)

본 연구에서 사용한 응시 추적 기술은 입력 영상에서 찾은 얼굴의 특징 파라메타(두 눈과 입의 끝점)로부터 가상의 얼굴 평면(Virtual Facial Plane)의 방향을 통해 실제얼굴의 응시 방향을 추정하는 방식이다. 그림 3.은 응시 추적 기술의 전체 흐름도이다.

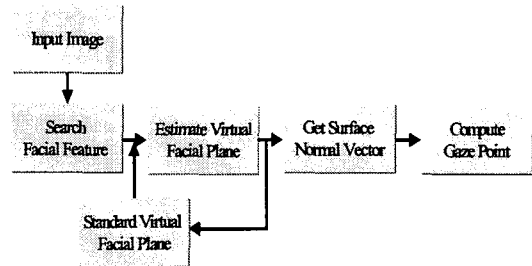


그림 3. Block Diagram of Gaze Tracking

1. 가상의 얼굴 평면(Virtual Facial Plane)

얼굴 평면(Facial Plane)은 얼굴의 특징점이 만드는 평면을 말한다[7]. 본 논문에서는 실제 얼굴의 특징 점들을 가지고 얼굴 평면을 설정하지 않고 그림 4.와 같이 초점(Pinhole)에서 실제 얼굴 평면 사이에 가상의 평면을 얼굴 평면으로 대체하였다.

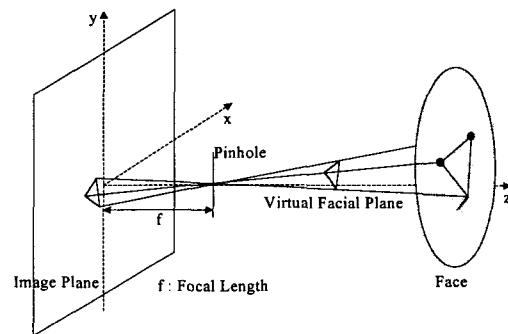


그림 4. Perspective Projection of Face

입력 영상에서 찾은 세 개의 특징 파라메타는 이미지 평면상에 투영된 실제 얼굴의 상(像)이기 때문에 이미지 평면상의 특징 파라메타에서 핀홀까지의 직선을 연장하여 그 세 개의 직선상의 점들을 추적해 가면 가상의 얼굴평면의 위치를 찾을 수 있다.

가. 기준 얼굴 평면(Standard Virtual Facial Plane) 설정
본 논문에서는 정면을 바라봤을 때, 카메라의 앞쪽으로 초점거리만큼 떨어진 위치에서 기준 얼굴 평면(Standard Virtual Facial Plane)을 설정하였다.

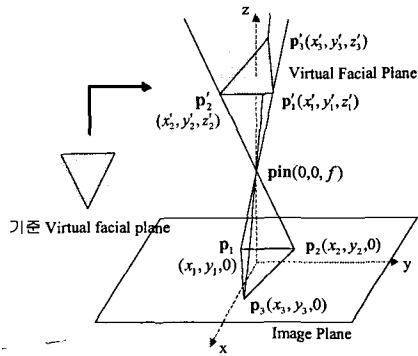


그림 5. Search of Virtual Facial Plane with Changed Pose

나. 변화된 위치의 얼굴 평면(Virtual Facial Plane)
그림 5.는 기준 얼굴 평면으로부터 변화된 위치에서의 얼굴 평면을 찾는 방법을 묘사하였다. 그림 5.와 같이 P_1', P_2', P_3' 는 가상의 얼굴 평면을 이루는 세 점으로, P_1, P_2, P_3 (이미지 평면상의 특징 파라메타)에서 pin (Pin hole)이 형성하는 직선 위에 존재하며 기준 얼굴 평면과 같은 크기를 갖는다. 단지 얼굴의 방향이 달라졌기 때문에 얼굴 평면의 위치(Pose)가 변화하였으므로 기준 얼굴 평면의 크기와 P_1, P_2, P_3 에서 pin 이 형성하는 직선의 방정식을 통해 P_1', P_2', P_3' 를 구할 수 있다. 이때 P_1', P_2', P_3' 의 위치를 추정하기 위해 기준 얼굴 평면이 세 직선과 접하도록 최대 경사법(Steepest Descent Method)을 이용해 에러율을 줄이는 방향으로 가상의 얼굴 평면(Virtual Facial Plane)의 세 점의 위치를 추정하였다.

2. 법선 벡터(Surface Normal Vector)
가상의 얼굴 평면의 법선 벡터는 평면 위의 두 벡터의 외적(Cross Product)을 통해 쉽게 구할 수 있다.

3. 응시위치(Gaze Point)
응시점은 가상의 얼굴 평면(Virtual Facial Plane)의 법선 벡터(Surface Normal Vector)와 이미지 평면이 만나는 점으로 계산될 수 있다.

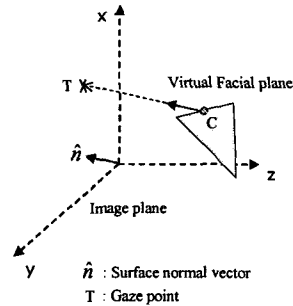


그림 6. Surface Normal Vector of Virtual Facial Plane

그림 6.은 얼굴 평면의 법선 벡터를 나타냈다. 여기서 얼굴 평면의 법선 벡터와 이미지 평면이 만나는 점이 T가 되고 이 점이 찾고자하는 응시점이 된다. 응시점을 구하기 위해 다음과 같은 식을 사용하였다.

$$\frac{\hat{n} - C}{\|\hat{n} - C\|} \times \|T - C\| + C = T \quad (1)$$

여기서 $\hat{n}(n_x, n_y, n_z)$ 을 얼굴 평면의 법선 벡터, $C(c_x, c_y, c_z)$ 를 얼굴 평면 상의 한 점, $T(T_x, T_y, T_z)$ 를 응시점이라고 할 때 식 (1)과 같이 응시점을 구할 수 있다. 여기서 이미지 평면의 정면에서 이미지 평면과 만나는 점을 찾고자 하므로 항상 $T_z=0$ 이 된다.

IV. 실험 결과 및 결론

응시 추적 시스템의 실험은 17인치 모니터를 3X3으로 9등분하여 각 블록의 중심점을 바라 봤을 때 프로그램 실행할 때 나타나는 화면(Window) 위에 모니터의 응시 위치와 동일한 부분에서 파란 색상이 나타나는 것으로 실시되었다. 실험환경은 USB로 인터페이스(Interface)되는 PC 카메라를 사용했으며, 펜티엄III (Pentium III) 1GHz CPU의 컴퓨터를 이용하였다. 응시 위치를 찾기 위해 모니터의 정면을 바라봤을 때의 얼굴 이미지를 기준 얼굴 평면으로 잡고 응시 위치를 변화해 가며 실험을 실시하였다.

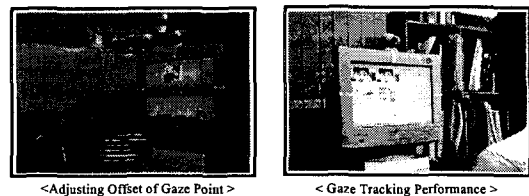


그림 7. Gaze Tracking System

실험에서 사용한 이미지의 크기는 352X288 (pixel)이고 초당 15프레임(frame)의 비디오 영상을 초당 5frame 의 이미지를 캡처(Capture)한 후 응시 추적하였다.

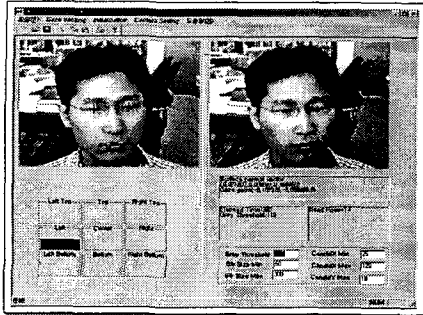


그림 8. Gaze Tracking Program

그림 9와 그림 10은 얼굴모델이 모니터의 9개의 지점을 바라봤을 때 측정한 이미지 평면상의 좌표값이다. 그림 9에서는 위(오른쪽 위, 왼쪽 위, 정면 위)와 아래(오른쪽 아래, 왼쪽 아래, 정면 아래)를 봤을 때 가상의 얼굴 평면의 법선벡터가 이미지 평면과 만나는 점의 Y값을 나타냈고, 그림 10은 왼쪽(왼쪽 위, 왼쪽 아래, 왼쪽)과 오른쪽(오른쪽 위, 오른쪽 아래, 오른쪽)을 봤을 때 이미지 평면상의 X값을 나타냈다. 위, 아래는 60개의 데이터 중 3개의 에러를 보여 5%의 에러율을 보였고 좌우는 18.3%의 에러율을 보였다.

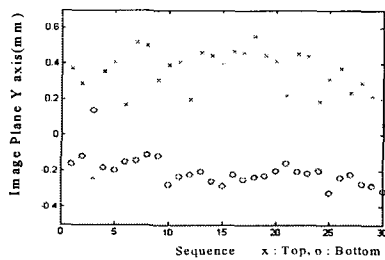


그림 9. 상하의 응시방향

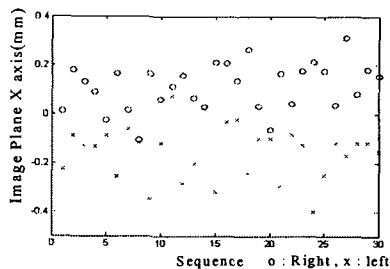


그림 10. 좌우의 응시방향

PC 카메라의 자동 초점기능이나 자동 노출 기능이 0.2초 정도의 영상 처리에 빠른 대응을 하지 못해 이미지가 흐려지는 점을 나타냈다. 이것은 빛의 변화나 사물의 움직임에 따라 카메라의 적응속도가 느려지면서 나타나는 현상으로 실시간으로 데이터를 처리해야 하는 상황에서 얼굴 영상의 특징 점을 찾지 못하여 에러를 발생시키는 주요 원인으로 나타났다.

참 고 문 헌

- [1] Jong-Gook Ko, "Facial Feature Tracking and Head Orientation based- Gaze Tracking", Department of Information and Communications, KJIST, Thesis for Master's Degree 2000.
- [2] A.L. Yuille, D.S. Cohen and P.W.Halnan, "Feature Extraction from Faces using Deformable Template", Proc. IEEE Computer Soc. Conf. on Computer Vision and Patt. Recog., pp. 104-109, 1989
- [3] Kyung-Nam Kim "Contributions to Vision-Based Eye-Gaze Tracking ", Department of Information and Communications, KJIST, Thesis for Master's Degree 1999.
- [4] S.H. Jeng, H.Y. Mark Liao, C.C. Han, M.Y. Chern and Y.T.Liu, "Facial Feature Detection Using Geometrical Face Model: An Efficient Approach", Pattern Recognition, Vol. 3 pp. 273-282, 1998.
- [5] I.Craw, H.Ellis, and J.Lishmana, "Automatic Extraction of Face Features", Pattern Recognition Lett., Vol 5, pp. 183-187, 1987.
- [6] "다차원 감성입력 기능을 갖는 교통정보안내 및 자동항법 장치", 과학기술부, 2000
- [7] Andrew Gee and Robergo Cipolla, "Determining the Gaze of Faces in Images", University of Cambridge, In Image and Vision Computing, Vol. 12, 1994
- [8] Arie E. Kaufman, Amit Bandopadhyay, Bernard D. Shaviv, "An Eye Tracking Computer User Interface", Research Frontier in Virtual Reality Workshop Proceedings, IEEE Computer Society Press, pp. 78-84, October 1993
- [9] 2단계 G7 감성 공학 기술 개발 사업 "컴퓨터의 감성 인터페이스 기술 개발", 연세대학교