

## 수정 Starburst 알고리즘과 Homography Normalization을 이용한 시선추적

조태훈\* · 강현민

### Gaze Tracking Using a Modified Starburst Algorithm and Homography Normalization

Tai-Hoon Cho\* · Hyun-Min Kang

School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan 330-708, Korea

#### 요 약

본 논문에서는 두 개의 카메라를 이용하여 보다 정확한 동공 인식을 통한 원격방식의 시선 추적을 제안한다. 헤드 장착형 시선추적용으로 개발된 Starburst 알고리즘은 원격방식의 시선추적에서는 카메라가 보다 넓은 영역을 보기 때문에 눈썹, 눈꼬리 등 외란이 많아 스타버스트 알고리즘을 바로 적용하면 동공 중심 추출에 실패하는 경우가 많았다. 이에 템플릿매칭을 이용하여 대략적인 동공영역을 찾고, 찾은 영역 내에서만 스타버스트 알고리즘으로 동공의 경계 후보점들을 찾은 후 보완된 RANSAC 알고리즘으로 타원근사하여 동공의 중심을 추출하였다. 추출된 동공 중심을 머리의 움직임에 거의 영향을 받지 않도록 4개의 적외선 LED를 모니터 네 구석에 부착하고 Homography normalization을 적용하였다. 스크린 좌표계로 변환할 때 기존에는 호모그래피를 사용하였으나, 카메라 렌즈의 비선형왜곡을 보상하기 위해 여기서는 고차다항식을 이용한 캘리브레이션 기법을 이용하였다. 끝으로, 두 대의 카메라를 사용하여 정확도와 신뢰성이 향상됨을 보인다.

#### ABSTRACT

In this paper, an accurate remote gaze tracking method with two cameras is presented using a modified Starburst algorithm and homography normalization. Starburst algorithm, which was originally developed for head-mounted systems, often fails in detecting accurate pupil centers in remote tracking systems with a larger field of view due to lots of noises. A region of interest area for pupil is found using template matching, and then only within this area Starburst algorithm is applied to yield pupil boundary candidate points. These are used in improved RANSAC ellipse fitting to produce the pupil center. For gaze estimation robust to head movement, an improved homography normalization using four LEDs and calibration based on high order polynomials is proposed. Finally, it is shown that accuracy and robustness of the system is improved using two cameras rather than one camera.

**키워드** : 스타버스트 알고리즘, 호모그래피, 시선 추적, 동공 인식

**Key word** : Starburst algorithm, homography normalization, gaze tracking, pupil detection

접수일자 : 2014. 02. 20 심사완료일자 : 2014. 03. 25 게재확정일자 : 2014. 04. 11

\* Corresponding Author Tai-Hoon Cho (E-mail: thcho@koreatech.ac.kr, Tel: +82-41-560-1351)

School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan 330-708, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2014.18.5.1162>

print ISSN: 2234-4772 online ISSN: 2288-4165

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.  
Copyright © The Korea Institute of Information and Communication Engineering.

## I. 서 론

시선 추적(gaze tracking)은 눈의 위치를 인식하고 눈이 응시하는 방향을 추정하고 추적하는 것을 말한다. 건강한 시선 추적은 손을 사용하지 못하는 사람들이 안구의 움직임으로 커서를 움직이게 하는 안구 마우스, 주의(attention) 분석 등의 많은 응용분야에서 매우 중요한 역할을 한다. 점점 저가, 고성능의 카메라들이 나오고, 컴퓨터 성능이 비약적으로 향상됨에 따라 비디오 기반 시선추적시스템들이 많이 연구 개발되고 있다[1].

시선 추정 및 추적은 크게 두 가지 부분, 영상 내에서의 눈 위치 검출 및 시선 추정으로 구성되며, 이 과정이 매 영상프레임 반복되어 시선 추적이 이루어진다. 눈의 위치는 흔히 동공 또는 홍채의 중심으로 구하는데, 그림 1과 같이 홍채는 영상에서 윤곽이 가려지는 경우가 많이 발생하기 때문에 홍채 안쪽의 작은 동공의 중심을 구하는 것이 신뢰성이 높다[2]. 이때, 적외선 카메라 영상이 가시광선대역 카메라 영상보다 동공과 홍채와의 경계가 훨씬 뚜렷하여 적외선 카메라를 많이 사용한다.

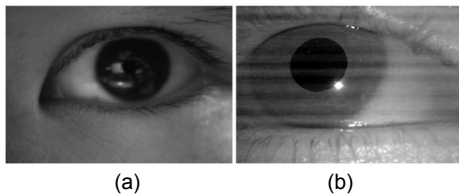


그림 1. (a) 가시광선 영상 (b) 적외선 영상  
Fig. 1 (a) Visible spectrum imaging (b) infrared spectrum imaging [2]

시선 추적 시스템은 카메라의 장착 위치에 따라 보통 카메라를 머리에 부착하여 카메라가 바로 눈앞에서 찍게 하는 헤드 마운트(Head-mounted)방식과 카메라를 모니터 앞 아래에 놓고 찍는 리모트(Remote)방식으로 나뉜다. 헤드 마운트 방식은 카메라가 눈앞 가까이에서 찍기 때문에 카메라의 시야범위를 눈 부근 영역만으로 좁게 한정할 수 있어 영상해상도가 높아져서 보다 정확한 눈 중심 추출이 가능하고, 머리 움직임에 영향을 받지 않는 장점이 있지만 사람이 항상 장비를 머리에 착용하고 있어야 하는 불편함이 있다. 반면에 리모트 방식은 카메라가 보통 모니터 밑에 고정되어 있어 어느 정도의 머리 움직임을 허용하려면 카메라의 시야범

위가 헤드마운트 방식보다는 좀 더 넓어야 하므로 영상 해상도가 낮아지고, 머리 움직임에 영향을 받기 때문에 눈 중심 추출 정확도가 상대적으로 떨어지지만, 장비를 머리에 착용하지 않아도 되는 장점이 있다.

헤드 마운트 방식으로 눈 검출 및 시선 추적하는 시스템으로 대표적인 예는 Starburst 알고리즘[2, 3]으로 공개 소프트웨어로 제공되어 많이 사용되고 있다. 하지만, 이것은 헤드마운트 방식을 가정하고 구현된 알고리즘이라서 카메라의 시야범위가 좀 더 넓은 리모트 방식에 바로 적용할 경우에 눈 이외의 다른 부분(속눈썹, 눈썹, 눈꼬리, 눈 주위 피부 등)에 의한 외란으로 눈 중심 검출이 제대로 되지 않는 경우가 자주 발생한다.

리모트 시선 추적 방식으로 사용상의 편의성 때문에 최근 많은 연구가 이루어져 상용화된 제품(예로, [4])도 있다. 리모트 방식의 가장 큰 문제는 사용자의 머리의 움직임을 효과적으로 보상할 수 있는 방법이다. 이를 위해 흔히 두 개 이상의 점광원(보통, 적외선 LED)을 모니터 스크린 테두리에 부착하는 방법을 많이 사용한다. 점광원이 각막에 반사되어 나타나는 점과 동공의 중심으로부터 시선의 방향을 눈의 기하학적인 3D 모델에 기반하여 추정한다. 이를 위해 일반적으로 여러 가지의 눈 모델, 카메라 파라미터를 구하기 위한 복잡한 캘리브레이션 과정이 요구된다[5].

이러한 복잡한 캘리브레이션 과정을 보다 단순화한 여러 방법들이 제시되었는데, 그 중 하나는 한 개의 카메라와 모니터 모서리에 부착된 네 개의 적외선 LED를 사용하고 Homography Normalization을 이용하는 것이다[6]. 4개의 적외선 LED에서 나온 빛은 각막(cornea)에 4개의 반사점으로 나타나고 이를 카메라에서 영상으로 취득하게 된다. 이때, 각막의 표면은 곡면이지만 카메라/스크린과의 거리가 훨씬 크기 때문에 각막의 네 점은 동일 평면에 있다고 근사화할 수 있다. 따라서 카메라 평면과 스크린 평면 사이에 각막평면을 경유한 호모그래피 매핑(mapping)이 존재한다고 볼 수 있다. 영상 좌표계에서 각 적외선 LED 반사점들( $g_1^i, g_2^i, g_3^i, g_4^i$ )을 길이가 1인 정사각형인 새로운 좌표( $g_1^n, g_2^n, g_3^n, g_4^n$ )의 노멀라이즈(Normalized) 좌표계로 바꿀 수 있는 호모그래피 매핑  $H_t^n$ 를 구하면, 이를 이용하여 이미지 좌표계에서 구한 동공의 중심  $p_c^i$ 로부터 노멀라이즈 좌표계 안의 대응되는 점  $p_c^n$ 을 구할 수 있다. 그리고, 노멀라이즈 좌

표계를 스크린좌표계로 매핑하는 행렬  $F_n^s$ 를 구하게 되면 노멀라이즈 좌표계의  $p_c^n$ 를 스크린 좌표계로 맵핑시킬 수 있다(그림 2)  $F_n^s$ 는 여러 개의 미리 정해진 스크린 좌표를 사용자가 순차적으로 응시하는 캘리브레이션 과정을 거쳐 호모그래피 매핑으로 구해진다[6].

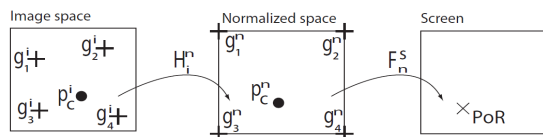


그림 2. 영상좌표계에서 노멀라이즈 좌표계, 그리고 스크린으로의 매핑 [6]

Fig. 2 Mapping from image space to normalized space and Screen [6]

본 논문에서는 개선된 리모트 방식의 시선 추적 시스템을 제안한다. 동공 중심 추출을 위해서는 헤드마운트 방식의 스타버스트 알고리즘을 리모트 방식에 맞도록, 수정, 보완하여 사용하였고, 머리의 움직임에 강인한 시선추적을 위해 [6]과 같이 네 개의 적외선 LED를 사용한 Homography Normalization을 이용하였다. 보다 정확한 카메라 렌즈의 비선형왜곡 보정을 위해  $F_n^s$ 를 구할 때 호모그래피를 사용하지 않고 고차다항식을 사용하였다. 또한 두 개의 카메라를 사용하여 시선추적의 정확성과 신뢰성을 향상시켰다.

## II. 시스템 구성 및 알고리즘

하드웨어는 크게 카메라, 렌즈, 적외선 필터, 적외선 LED, 모니터로 이루어져 있다. 카메라는 근적외선 영역에서 감도가 양호한 PointGrey사의 Firefly MV (FFMV-03) 두 대(USB2, IEEE1394)를 사용하였다. 이 센서는 1/3" progressive scan CMOS 센서로 해상도는 752 x 480, 최대 프레임 레이트는 60장/초이다. 카메라는 그림 3과 같이 모니터의 하단부에 2대의 카메라가 각각 왼쪽과 오른쪽 눈을 찍을 수 있도록 위치시켰다. 렌즈는 초점 거리가 50mm인 CCTV 렌즈를 사용하였고 렌즈 앞에는 가시광선을 차단하는 적외선 통과 필터를 장착하였다. 적외선 LED는 모두 4개를 사용하였는데, 모니터의 네 모서리에 하나씩 부착하고 LED가 사

람의 눈 쪽을 향하도록 하였다. 모니터는 1280x1024 해상도의 19인치 모니터를 사용하였으며 사용자와 카메라의 거리는 약 60~70 센티 정도이다.

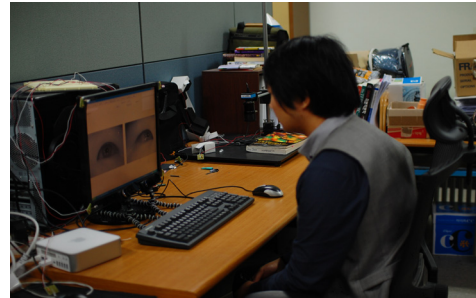


그림 3. 시스템 셋업

Fig. 3 System setup

전체적인 알고리즘 흐름도는 그림 4와 같다.

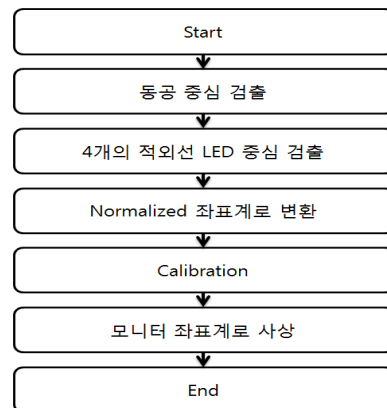


그림 4. 전체적인 알고리즘 흐름도

Fig. 4 Overall flow of the algorithm

### 2.1. 동공 중심 추출

동공 중심 추출은 그림 5와 같이 이루어진다. 동공의 중심을 찾는데 있어 4개의 적외선 LED의 각막 반사(corneal reflection)는 잡음으로 작용한다. 그래서 동공의 중심을 추출하기 전에 4개의 적외선 LED의 반사점을 찾아서 제거할 필요가 있다. 먼저, 이진화에 의해서 가장 밝은 부분을 검출한다. 이 부분은 적외선 LED의 반사점과 피부반사에 의한 잡음, 그리고 눈의 공막(sclera)부분의 반사에 의한 부분으로, 모두 적외선 LED 반사점의 후보군으로 선택한다.

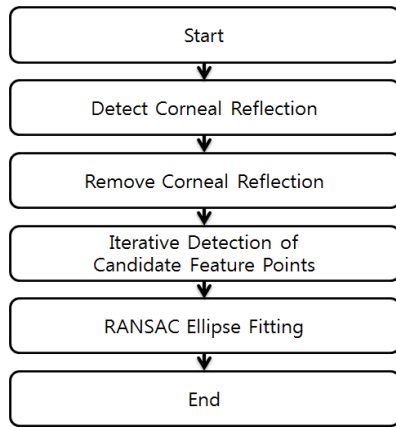


그림 5. 동공 중심 추출 알고리즘 순서도  
Fig. 5 Flow of pupil center extraction

적외선 LED의 반사점을 제거하기 위해서 처음엔 Starburst 소스에 있는 알고리즘을 적용해 보았지만 완벽하게 동작하지 않아, 좀 더 간단하지만 확실한 방법을 사용하였다. 위에서 찾은 **적외선 LED의 반사점의 외곽선을 두꺼운 선으로 그리는 방법을 사용하였다**. 영상에서 실제 적외선 LED의 반사점은 몇 픽셀이 되지 않기 때문에 조금만 두꺼운 선으로 그려도 충분히 반사점을 제거할 수 있다. 물론 이렇게 하였을 경우 자연스럽게 제거가 되지 않는다. 제거한 부분이 잡음으로 나올 수 있지만 후 보정을 통해 이러한 잡음들은 제거할 수 있다. 기존의 Starburst는 헤드 마운트 방식을 사용하였기 때문에 카메라의 시야 범위가 작아 눈 부분만 나온다. 하지만 본 논문에서는 헤드 마운트 방식이 아닌 리모트 방식을 사용하기 때문에 그림 6과 같이 이미지의 영역이 눈만 나오지 않고 눈썹 등 외란이 많다. 따라서 스타버스트 알고리즘으로 전체 영상 영역에서 동공을 찾으려면 경우에 따라 제대로 찾을 수도 있지만 눈썹 등과 같은 부분에 영향을 받아 동공 경계점 후보군을 제대로 찾지 못할 때도 있다. 때문에 스타버스트 알고리즘을 적용할 영역을 최대한 줄이기 위해, **LED 반사점이 없는 동공 영역을 템플릿으로 하여 동공의 대략적인 위치를 찾아, 템플릿의 영역 크기 만큼 관심영역을 설정한 후, 이 관심영역 내에서만 스타버스트 알고리즘을 적용하면 거의 동공 특징점들만 추출할 수 있다. 동공의 중심을 찾기 위해서 동공 경계 특징점들을 찾지만 잡음 또한 검출하게 된다. 이러한 잡음들에 대해 강인하게 동공의 중심을 찾기 위해 RANSAC (Random SAmple Consensus)**

알고리즘 [7]을 사용한다. 그림 7은 RANSAC을 이용한 타원근사 알고리즘의 순서도이다.

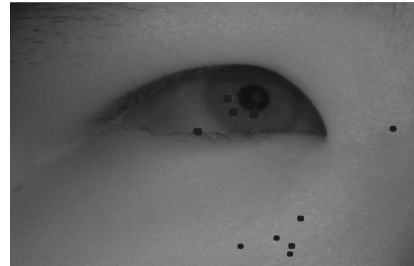


그림 6. 적외선 LED 반사점과 잡음들  
Fig. 6 Reflection points of infrared LEDs and noises

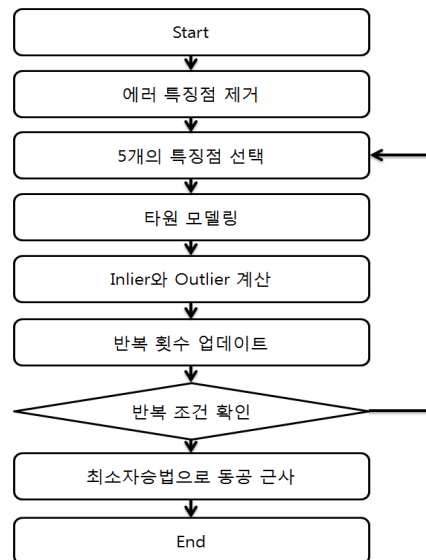


그림 7. RANSAC 동공 타원 피팅 알고리즘  
Fig. 7 Ellipse fitting of pupil boundary feature points extracted using RANSAC

Starburst 알고리즘으로 찾을 때 동공 이외의 부분과 적외선 LED 제거를 할 때 윤곽을 두꺼운 선으로 칠해서 제거했기 때문에 이 부분에서 잘못된 특징점들을 찾을 수가 있다. 따라서 이러한 부분들의 특징점들을 제거한다. 먼저, 템플릿 매칭의 관심 영역 안에 없는 특징점과 적외선 LED 후보군의 경계에 있는 특징점들을 제거한다. 타원으로 피팅을 하기 위해서는 최소한 5개의 점이 필요하다. 찾은 특징 점 후보군들 중에서 임의의 5 점을 선택한다. 임의의 점 5개를 선택하면 타원의 방정

식을 구할 수 있다. 일반적인 타원의 방정식은 식(1)과 같다.

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (1)$$

식(1)을 타원의 중심이  $(c_x, c_y)$ 인 표준형태의 타원 식으로 변환을 하게 되면 식(2)가 된다.

$$\frac{(\cos\theta(x-c_x) - \sin\theta(y-c_y))^2}{w^2} + \frac{(\sin\theta(x-c_x) + \cos\theta(y-c_y))^2}{h^2} = 1 \quad (2)$$

여기서, 타원 회전 각도

$$\theta = \frac{1}{2} \tan^{-1} \left( \frac{b}{a-c} \right) \quad (3)$$

$$c_x = \frac{2cd - be}{b^2 - 4ac} \quad (4)$$

$$c_y = \frac{2ae - bd}{b^2 - 4ac} \quad (5)$$

$$w = \sqrt{\frac{ac_x^2 + bc_xc_y + cc_y^2 - f}{a\cos^2\theta + b\cos\theta\sin\theta + c\sin^2\theta}} \quad (6)$$

$$h = \sqrt{\frac{ac_x^2 + bc_xc_y + cc_y^2 - f}{a\sin^2\theta - b\cos\theta\sin\theta + c\cos^2\theta}} \quad (7)$$

스타버스트 알고리즘에서는 인라이어(inlier)를 판단할 때 식 (1)의 왼쪽의 대수거리를 사용한다. 일반적으로 대수거리보다는 기하학적인 거리를 사용하는 것이 보다 강건한 타원 피팅을 이룰 수 있다. 여기에서는 기하학적인 거리를 이용하여 인라이어를 판단한다.

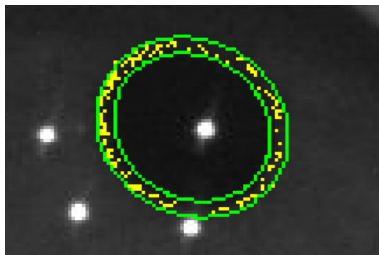


그림 8. 인라이어 결정  
Fig. 8 Determination of inlier points

식 (2)를 이용하여 장축과 단축이 일정 임계값 만큼 작은 타원과 큰 타원을 구해 작은 타원과 큰 타원 안에 있는 특징점은 인라이어(inlier)로, 바깥에 있는 특징점을 아웃라이어(outlier)로 판별한다. 그림 8은 위 식으로 구한 타원으로 장축과 단축의 임계값을  $\pm 10\%$ 로 하여 구한 작은 타원과 큰 타원 그리고 그 안에 들어 있는 인라이어를 표시한 그림이다.

RANSAC의 반복횟수  $N$ 은 확률  $p$ (보통 0.99로 설정)를 보장할 수 있도록 충분히 커야 한다. 확률  $p$ 는 최소한 하나의 샘플 집합이 유효한 데이터(inlier)만을 포함할 확률이다. 선택된 데이터가 인라이어일 확률을  $w$ 라고 하면  $N$ 은 다음과 같이 구한다. 여기서  $s$ 는 한번 샘플할 때 필요한 데이터의 개수로 여기서는 5이다.

$$1 - p = (1 - w^s)^N \quad (8)$$

$$N = \frac{\log(1 - p)}{\log(1 - w^s)} \quad (9)$$

현재 알고리즘의 반복 횟수가 최대 반복 횟수에 도달할 때까지 5개의 특징점을 다시 선택하여 처리하는 과정을 반복하게 된다. 최대 반복 횟수에 도달하면 구해진 인라이어들을 이용하여 최소자승법(Least squares method)으로 타원 피팅한다. 타원의 중심은 바로 동공의 중심이 된다.

## 2.2. 적외선 LED 중심 추출

위의 동공중심 검출과정에서 적외선 LED 반사점 후보군을 검출하였다. 하지만 이렇게 구하였을 때 동공의 외곽 쪽에 다른 반사에 의해 잡음이 발생할 수 있다. 그렇기 때문에 이런 잡음과 적외선 LED의 반사점을 구분할 필요가 있다. 잡음들이 생기는 부분은 대부분 동공의 바깥 부분에 주로 생긴다. 그렇기 때문에 Starburst 알고리즘으로 구한 동공의 중심으로부터 거리를 이용하여 적외선 LED 반사점의 위치와 잡음을 구분할 수 있다.

먼저 템플릿 매칭을 해서 나온 관심영역 안에 들어 있는 적외선 LED 반사점 후보군들만 선택을 한다. 적외선 LED 반사점의 후보군들이 4개보다 많을 경우 동공의 중심까지의 거리를 구해 그중 가장 가까운 4개의 포인트를 LED 반사점으로 선택한다. 적외선 LED 반사점의 중심은 이진화 영상에서 각각의 영역(Blob)별로 무게 중심으로 구한다[8].



### 2.3. Homography Normalization

호모그래피(Homography)는 하나의 평면을 다른 하나의 평면으로 옮기는 투영 변화를 의미한다. 2개의 이미지에 서로 대응되는 한 점  $(x_i, y_i)$ 와  $(\bar{x}_i, \bar{y}_i)$ 가 있다면 두 이미지에 관련된 호모그래피 H는 다음과 같다.

$$\begin{bmatrix} w\bar{x}_i \\ w\bar{y}_i \\ w \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (10)$$

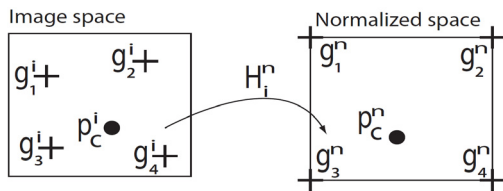


그림 9. 영상좌표계에서 노멀라이즈좌표계로의 호모그래피  
Fig. 9 Homography from image space to normalized space

4개의 적외선 LED를 모니터의 모서리에 붙였고 사람의 시선이 모니터를 응시한다면 항상 4개의 적외선 LED가 만들게 되는 사각형 안에 존재한다.

각각의 적외선 LED 반사점들( $g_1^i, g_2^i, g_3^i, g_4^i$ )과 새로운 좌표계( $g_1^n, g_2^n, g_3^n, g_4^n$ )인 노멀라이즈 좌표계로 바꿀 수 있는 호모그래피 행렬 H를 구하면 영상(Image space)안의 동공의 중심  $p_c^i$ 가 노멀라이즈 좌표계 안에 대응하게 되는 점  $p_c^n$ 을 구할 수 있다(그림 9).

### 2.4. 캘리브레이션

시선의 위치를 구하기 위해, 노멀라이즈 좌표계에서 구한 동공의 중심은 스크린 좌표계로 변환하는 매핑 함수를 구하는 캘리브레이션 단계가 필요하다. 캘리브레이션은 보통 그림 10과 같이 등간격으로 분포되어 있는 캘리브레이션 점들을 갖고 있는 캘리브레이션 화면을 모니터에 전체 화면으로 띄우고 사용자가 각 점을 순차적으로 바라보면서 이루어진다. 각 점을 바라볼 때 동공의 중심을 추출하여 노멀라이즈 좌표계로 변환되고, 이 좌표는 응시하고 있는 캘리브레이션 화면 스크린 좌표에 대응된다(그림 11). (캘리브레이션 점 스크린 좌표는 이미 알고 있음)

N개의 캘리브레이션 점 좌표  $(x_i, y_i)$ 가 있다면, 동공

중심 추출 과정을 통하여 구해진 N개의 동공중심 노멀라이즈된 좌표  $(I_i, J_i)$ 가 대응된다.

이를 이용하여, 동공중심 좌표를 스크린 좌표로 매핑하는 함수를 구하게 된다. [6]에서는 이를 위해 호모그래피 매핑을 사용하였으나, 호모그래피 매핑은 카메라 렌즈의 비선형 왜곡에 대한 보정이 거의 어렵다. 그래서 여기서는 카메라렌즈의 비선형 왜곡의 보정이 가능한 고차 다항식에 의한 매핑[9]을 사용하여 캘리브레이션을 수행하였다.

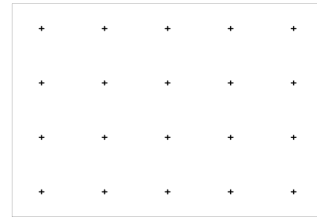


그림 10. 캘리브레이션 점  
Fig. 10 Calibration points

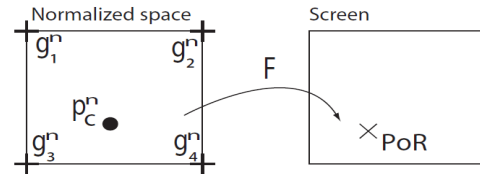


그림 11. 노멀라이즈 좌표계에서 스크린으로의 매핑  
Fig. 11 Mapping from normalized space to screen

노멀라이즈 좌표  $(I, J)$ 와 캘리브레이션 점 스크린 좌표  $(x, y)$ 와의 관계를 M차 다항식 변환으로 표시하면 아래와 같다.

$$x = \sum_{m=0}^M \sum_{n=0}^M k_{mn} I^m J^n, y = \sum_{m=0}^M \sum_{n=0}^M l_{mn} I^m J^n \quad (11)$$

N개의 캘리브레이션 좌표와 이에 대응되는 노멀라이즈 좌표가 주어졌을 때, 위식을 최소자승오차 (Least Square Error)상 가장 정확하게 만족시키는

즉,  $\sum_{i=1}^N (x_i - \sum_{m=0}^M \sum_{n=0}^M k_{mn} I^m J^n)^2$ 을 최소화하는  $k_{mn}$ ,  $\sum_{i=1}^N (y_i - \sum_{m=0}^M \sum_{n=0}^M l_{mn} I^m J^n)^2$ 을 최소화하는  $l_{mn}$ ,  $0 \leq m, n \leq M$ 을 구한다.

$N = (M+1)^2$ 이면 유일해(Unique Solution)가 존재하고,  $N > (M+1)^2$ 이면 overdetermined 선형방정식이 되어 pseudo-inverse 방법에 의해 해를 구한다.

보통 2차나 3차 다항식을 많이 사용하는데, 2차일 경우 최소 9개의 캘리브레이션 점이 필요하며, 3차인 경우에는 최소 16개의 점이 요구된다. 캘리브레이션 점이 많을수록 보다 정확한 매핑함수를 구할 수 있지만, 사용자의 불편도 따라 증가하게 된다. 여기서는 2차 다항식을 사용하고 20개의 캘리브레이션 점을 사용하였다.

캘리브레이션을 통하여 실제 화면 상의 좌표로 매핑되지만, 여러 가지 요인에 의한 잡음이 생겨 실제로 의도하지 않았던 곳으로 좌표가 매핑될 수 있다. 이런 경우 캘리브레이션을 했을 경우 오차가 크게 발생한다. 따라서, 캘리브레이션을 할 때 각 좌표마다 한 장의 이미지로 캘리브레이션을 하는 것 보다 각 좌표마다 여러 장의 이미지를 이용하여 구한 각 좌표들의 미디언(median) 값을 사용하면 잡음에 좀 더 영향을 받지 않고 캘리브레이션을 할 수 있다. 본 논문에서는 x축과 y축에서 각각 미디언 값을 취해 사용하였다.

### III. 실험 결과

실험 영상은 각각의 카메라가 왼쪽 눈과 오른쪽 눈을 각각 찍어서 한번에 30장/초로 30장씩 찍었다. 영상의 크기는 640x 480이다. 찍은 영상은 20개의 캘리브레이션 포인트를 각각 바라보는 왼쪽과 오른쪽 눈 영상이다.

왼쪽과 오른쪽 눈 영상을 이용하여 동공 중심 추출, 적외선 LED 중심 추출, 이미지 상의 좌표를 노멀라이즈 좌표계로 사상시킨다. 그리고 이것을 다시 스크린 좌표계로 사상시켜 왼쪽과 오른쪽 눈 좌표 그리고 평균을 이용한 좌표를 실제 캘리브레이션 좌표와 비교하였다. 캘리브레이션은 2차 다항식을 이용했으며 캘리브레이션때 5장의 미디언 값을 사용하였다.

그림 12는 캘리브레이션 각 좌표점에 대해 1장의 영상만을 사용해서 캘리브레이션을 했을 때 왼쪽과 오른쪽 x,y에 대한 RMS(Root Mean Square) 에러와 5장의 영상의 미디언 값을 사용했을 때 왼쪽과 오른쪽 x,y에 대한 RMS 에러값을 나타낸 것이다. 그림 12의 4번 그래프는 잡음이 생겨 왼쪽과 오른쪽 x의 RMS 값이 크게 나왔다. 이런 RMS를 가지고 캘리브레이션을 하게 되면

스크린 상에 실제로 원하는 곳을 사상시키기가 힘들다. 그러므로 1장이 아닌 여러 장의 영상을 사용해 미디언 값을 사용하면 이런 에러가 날 확률을 줄일 수 있다.

그림 13은 20개의 캘리브레이션 점을 스크린 좌표계로 사상시켰을 때 각도 에러(단위: 도)이다. 시선(gaze) 추정 오차는 보통 각도로 나타내는데, 현재 상용되고 있는 제품의 보통 오차는 보통 1°정도 이다. 사람의 눈과 모니터의 거리를 r, 시선추정 각도 오차를  $\theta$ , 스크린 상에서 시선 위치 오차거리를 l이라고 했을 때  $\theta$ 가 매우 작으면,

$$l \cong r\theta \quad (12)$$

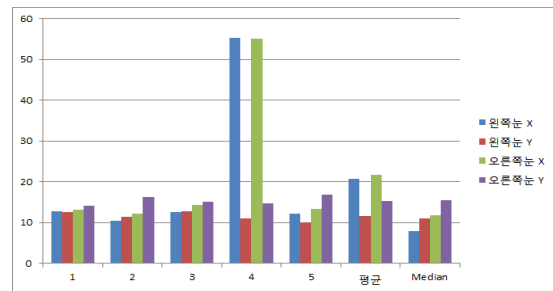


그림 12. 캘리브레이션 RMS 에러 [픽셀]  
Fig. 12 Calibration RMS error [pixel]

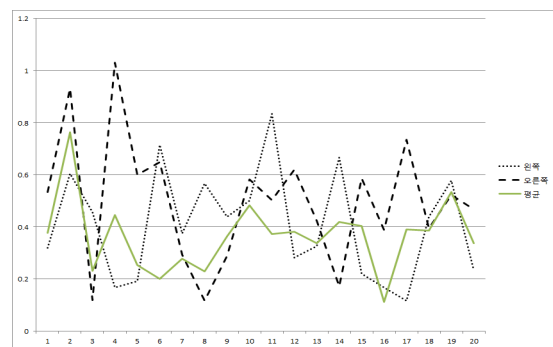


그림 13. 캘리브레이션 에러 [도]  
Fig. 13 Calibration error [degree]

사람의 눈과 모니터와 거리를 약 70cm, 모니터의 해상도는 1280 x 1024, 모니터 패널의 가로 길이는 38cm 일 때 계산을 해보면 최대 오차 범위는 약 37 Pixel일 때 약 1.03도의 시선 각도 오차를 갖는다.

왼쪽 카메라 영상만으로 캘리브레이션했을 때 최소 0.12, 최대 0.83, 평균 0.4도의 에러를 보이고, 오른쪽 영상만으로 캘리브레이션한 경우 최소 0.12, 최대 1.03, 평균 0.5도의 오차를 나타냈다. 왼쪽과 오른쪽에서 구한 스크린 추정좌표를 평균해서 구하면 최소 0.11, 최대 0.76, 평균 0.36도의 오차를 보였다. 두 대의 카메라를 사용하는 것이 하나의 카메라를 사용할 때 보다 최대 26% 정확도가 향상되는 것을 알 수 있다.

#### IV. 결 론

본 논문에서는 두 개의 카메라를 이용하여 보다 정확한 동공 인식을 통한 시선 추적을 제안하였다. 동공 중심 추출을 위해 Starburst 소스를 수정 보완하였다. 기존의 스타버스트 알고리즘은 헤드 장착형 시선추적용으로 개발되어 대부분 동공 및 홍채부분 만 영상으로 취득되어 스타버스트 알고리즘이 잘 동작하지만, 여기서 사용하는 원격방식의 시선추적에서는 카메라가 보다 넓은 영역을 보기 때문에 눈썹, 눈꼬리 등 외란이 많아 스타버스트 알고리즘을 바로 적용하면 동공 중심 추출에 실패하는 경우가 많았다. 이에 템플릿매칭을 이용하여 대략적인 동공영역을 찾고, 찾은 영역 내에서만 스타버스트 알고리즘으로 동공의 경계 후보점들을 찾은 후 보완된 RANSAC 알고리즘으로 타원근사하여 동공의 중심을 추출하였다.

추출된 동공중심을 머리의 움직임에 거의 영향을 받지 않고 스크린 좌표로 변환하기 위해 4개의 적외선 LED를 모니터 네 구석에 부착하고 Homography normalization을 적용하였다. 노멀라이즈 좌표계를 스크린 좌표계로 변환할 때 기존에는 호모그래피를 사용하였으나, 카메라 렌즈의 비선형왜곡을 보상하기 위해 여기서는 고차다항식을 이용한 캘리브레이션 기법을

이용하였다.

끝으로, 두 대의 카메라를 사용하는 것이 하나의 카메라를 사용할 때 보다 최대 26% 정확도가 향상되었다. 또한, 카메라를 두 개 사용하게 되면 한쪽이 잘 찾지 못하더라도 다른 한쪽을 사용하면 되므로 한 개를 사용할 때보다는 신뢰성이 향상된다.

#### REFERENCES

- [1] D. W. Hansen, "In the Eye of the Beholder : A survey of Models for Eyes and Gaze," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 478-500, 2010.
- [2] Dongheng Li, "Low-cost eye-tracking for human computer interaction," MS Thesis, Iowa State Univ., 2006.
- [3] D. Li, D. Winfield, D.J. Parkhurst, "Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches," *Proc. of the IEEE Vision for Human and Computer Interaction Workshop at CVPR 2005*, pp.1-8, 2005.
- [4] <http://www.eyetechds.com>.
- [5] E.D. Guestin and M. Eizenman, "General Theory of Remote Gaze Estimation Using the Pupil Center and Corneal Reflections," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1124.1133, 2006.
- [6] D. W. Hansen, J. S. Agustin, A. Villanueva, "Homography Normalization for robust Gaze Estimation in Uncalibrated Setups," *ETRA '10 Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 2010.
- [7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, 2003.
- [8] R. Jain, R. Kasturi, B. G. Schunck, *Machine Vision*, 1995.
- [9] T.-H. Cho, "Accurate camera calibration using high order polynomials," *Korean Institute of Intelligent Systems 2007 Spring Conference Proceedings*, vol. 17, no. 1, pp.413-416.



조태훈 (Tai-Hoon Cho)

1991. 5 : Virginia Tech 전기공학과 공학박사  
 1992. 3 ~ 1998. 2 : LG산전연구소 수석연구원  
 1998. 3 ~ 현재 : 한국기술교육대학교 컴퓨터공학부 교수  
 ※관심분야 : 영상처리, 컴퓨터비전





**강현민 (Hyun-Min Kang)**

2010. 2 : 한국기술교육대학교 정보기술공학부 학사

2012. 2 : 한국기술교육대학교 컴퓨터공학과 석사

※관심분야 : 영상처리, 컴퓨터비전