creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

[UCI]I804:24011-200000237582

2009년 2월
教育學碩士(情報·컴퓨터教育)學位論文

베이지안 네트워크에서 Kalman필터 알고리즘을 이용한 실시간 시선식별에 관한 연구

朝鮮大學校 教育大學院

情報 컴퓨터 教育 傳攻

尹 泳 智

베이지안 네트워크에서 Kalman필터 알고리즘을 이용한 실시간 시선식별에 관한 연구

A Study on real time Gaze Discimination Using Kalman Fillter Algorithm of Bayesian Network

2009年 2月

朝鮮大學校 敎育大學院

情報 컴퓨터 敎育 傳攻

尹 泳 智

베이지안 네트워크에서 Kalman필터 알고리즘을 이용한 실시간 시선식별에 관한 연구

指 導 敎 授　　羅　　相　　同

이 論文을 敎育學碩士(情報·컴퓨터 敎育專攻)學位

請求 論文으로 提出합니다.

2008年　10月

朝鮮大學校 敎育大學院

情報 컴퓨터 敎育 專攻

尹　　泳　　智

尹泳智의 敎育學 碩士學位 論文을 認准함.

審査委員長 朝鮮大學校 敎授 印

審査委員 朝鮮大學校 敎授 印

審査委員 朝鮮大學校 敎授 印

2008年 12月

朝鮮大學校 敎育大學院

목 차

I. 서 론 ····· 1

II. 베이지안 네트워크를 이용한 얼굴 특징추적 ····· 3
2.1 베이지안 네트워크 이론 ····· 3
2.2 눈동자 움직임을 이용한 시선 식별 법 ····· 8
2.3 얼굴 움직임을 이용한 시선 식별 법 ····· 13
2.4 얼굴 특징 추적 ····· 15
2.5 얼굴 동작 인식을 위한 인공 신경망 ····· 21

III. Kalman 필터 알고리즘 제안 ····· 24
3.1 Kalman 필터 이론 ····· 24
3.2 Kalman 필터 알고리즘 ····· 30
3.3 Kalman 필터와 평균이동 알고리즘을 이용한 동공 추적 ····· 32
3.4 시선 식별 전처리 ····· 38
3.5 매개변수 추출 ····· 45
3.6 시선교정 ····· 49

IV. 실험 결과 검증 및 평가 ····· 52
4.1 얼굴 동작 분류에 대한 시선 식별 ····· 52
4.2 학습에 참여하지 않은 사용자의 얼굴 움직임이 있는 실험 검증· 55
4.3 인접 영역에 대한 재식별 인식률 검증 ····· 57

V. 결　　론 ····· 61

참고문헌 ····· 63

Abstract ····· 68

그 림 목 차

그림 2-1 베이지안 네트워크 그래프 ····· 5
그림 2-2 눈동자내의 홍채, 동공 및 공막 영역 ····· 9
그림 2-3 얼굴의 3차원 움직임 량 추정 ····· 14
그림 2-4 얼굴의 구조적 특징 ····· 16
그림 2-5 얼굴 특징 요소 추출 ····· 18
그림 2-6 얼굴 영상에서 얼굴 영역 검출 ····· 20
그림 3-1 Kalman필터 루프 ····· 25
그림 3-2 Kalman필터 알고리즘 ····· 31
그림 3-3 동공 추적 알고리즘 ····· 34
그림 3-4 평균 이동 추적기를 이용한 추적 ····· 37
그림 3-5 밝고 어두운 동공 효과 원리 ····· 39
그림 3-6 하드웨어구성 : 능동적 적외선 조명 카메라 ····· 40
그림 3-7 (a)짝수필드영상 (b)홀수필드영상 (c)차 영상 ····· 43
그림 3-8 동공후보 영역을 표시한 영상 ····· 44
그림 3-9 시선 식별을 위한 매개변수 ····· 46
그림 3-10 얼굴방향 식별알고리즘 순서도 ····· 48
그림 3-11 시선교정을 위한 GRNN구조 ····· 50
그림 4-1 얼굴동작 분류 결과 ····· 53
그림 4-2 시선 영역에 대한 인접 영역 ····· 58
그림 4-3 재식별기를 이용한 시선 식별 결과 ····· 40

표 목 차

표 2-1 조건 확률 표 ····· 6

표 3-1 동공 추적 비교 ····· 36

표 4-1 학습에 참여하지 않은 사용자의 얼굴 움직임이 있는 상태에서의 시선식별 ····· 56

표 4-2 재식별기를 이용한 시선식별 결과 ····· 59

I. 서 론

멀티미디어 시스템을 활용한 시선 식별은 많은 응용 분야에서 이용되고 있다. 기존의 시선 식별 시스템은 시선을 식별하기 위해 얼굴의 움직임과 눈동자의 움직임을 이용했다[1][2]. 얼굴의 움직임만을 이용할 경우 얼굴의 위치에 따라 시선의 위치를 결정하게 되는데 미세한 시선 변화를 감지하지 못하는 단점이 있다. 현재까지 연구된 눈동자에 근거한 시선 추정 방식의 가장 큰 공통점은 각막의 글린트와 동공 사이의 상대적인 위치에 근거한다는 것이다[3][4]. 눈과 시선 식별 시스템의 또 다른 문제는 각각의 사용자에 대하여 교정 작업이 필요하다는 점이다.

그런데 가상현실 환경의 경우 사용자의 얼굴 움직임 및 시선 위치에 따라 삼차원 화면을 조정할 수 있다. 무인 안내 시스템에서 시선 식별 기술은 고가의 터치스크린 등을 대용할 수 있다. 특히, 사용자의 시선 식별을 위해서는 얼굴 및 눈, 코, 입 같은 얼굴 특징 점의 위치를 추적하는 부분이 필요하다. 이러한 기술은 무인 감시 시스템에서 얼굴 추적 및 인식으로 시선을 식별한다.

그리고 베이지안 네트워크에서 시선 식별은 현재 사용자가 응시하고 있는 위치를 자동적으로 식별해 내는 것으로 휴먼 컴퓨터 인터페이스의 한 연구 분야라고 할 수 있다. 휴먼 컴퓨터 인터페이스는 인간의 지능적 판단 능력을 모방하여 인간이 사용하는 의사소통 수단을 컴퓨터가 이해할 수 있도록 하는 혁신적인 기술이다[5]. 즉, 컴퓨터가 사람의 음성을 인식하고 그 명령을 처리함은 물론, 감성까지 이해하여 무엇을 보거나 듣거나 접촉할 때 갖는 감각이나 느낌을 컴퓨터가 이해한 후 적절하게 행동할 수 있도록 하는 것이 연구 목표이다.

제스처나 얼굴 표정, 음성, 시선 등을 이용한 멀티 모달(Multi-modal) 사용자 인터페이스에 대한 연구 중 마우스의 기능을 대신하여 사람의 시선 방향을 활용하는 방식은 컴퓨터가 사람의 의도를 인지한다는 측면에서 인간과 컴퓨터간의 새로운 의사 소통방식의 하나라고 할 수 있다[6].

본 논문에서는 기존 시선 식별 시스템의 문제점인 얼굴 움직임이 있을 시 시선

식별이 어려운 점과 사용자에 따른 교정 작업이 필요하다는 점을 해결하고자 베이지안 네트워크에서 Kalman 필터 알고리즘을 적용한 실시간 시선 식별 시스템을 제안한다.

시선 식별에서 영상 획득을 위하여 두 개의 고리 형태의 적외선 LED를 장치한 조명기와 그 중앙에 적외선 카메라를 장치한 영상 획득 도구를 구성한다. 영상 획득 시 짝수 필드와 홀수 필드에서 안쪽 LED와 바깥쪽 LED가 교차 점멸 하도록 하여 밝은 동공효과와 어두운 동공효과를 만들어 동공을 검출하고, 동공의 검출 후 객체 추적 알고리즘인 Kalman 필터 알고리즘에 의하여 연속되는 프레임에서 동공을 추적한다. 추적된 동공에서 글린트와 동공의 변화량과 동공의 타원정도, 동공과 글린트와의 기울기 정보 등을 추출한다.

II. 베이지안 네트워크에서 얼굴 특징 추적

2.1 베이지안 네트워크 이론

베이지안 네트워크는 특정 분야의 영역 지식을 확률적으로 표현하는 대표적인 수단으로 변수들 간의 확률적 의존 관계를 나타내는 그래프와 각 변수별 조건 확률들로 구성된다. 즉, 베이지안 네트워크는 각 노드(node)마다 하나의 조건 확률 표(CPT:Conditonal Probability table)를 갖는 비 순환 유향 그래프로 정의할 수 있다[9]. 베이지안 네트워크는 노드들에 배정된 값들을 기초로 목표노드가 가질 조건 확률을 계산하는데 이용할 수 있다.

따라서 베이지안 네트워크는 정보가 노드에 값이 주어졌을 때 숨김(hidden)노드의 확률 분포를 구하고 최종적으로 목표노드의 확률 값을 추론할 수 있다[7]. 즉, 베이지안 네트워크는 하나의 데이터 집합으로부터 학습할 때 각 노드는 데이터 집합의 각 특징을 표현하고 각 가지(arc)는 특징들 간의 의존성을 표현하게 되며 이렇게 학습된 베이지안 네트워크를 기초로 목표노드를 확률적으로 예측한다[8].

각 노드의 조건 확률과 노드 간의 연결선들로 이루어진 베이지안 네트워크 구조가 주어지면, 변수들에 대한 확률 분포 식은 식(2.1)과 같다.

$$p(x) = \prod_{i=1}^{n} p(x_i | pa_i) \quad \text{.....(2.1)}$$

pa_i는 그래프에서 x_i의 부모노드(parent node)의 집합을 나타낸다. 그래프에서 간선으로 연결되지 않은 노드들은 서로 조건부 독립 관계에 있다. 주어진 네트워크의 구조가 변수 집합 x의 물리적 확률 분포를 표현할 수 있다면 노드 x_i에 대해 영향을 주는 다른 모든 노드들과의 종속성은 식(2.2)와 같다.

$$p(x) = \prod_{i=1}^{n} p(x_i | x_1, \ldots\ldots x_{i-1}) \quad \ldots\ldots (2.2)$$

여기서, 변수의 집합을 x라 하면 $x = x_1, \ldots x_{i-1}$로 표시되고, x의 부모노드를 $p(x)$로 나타내면 식 (2.3)과 같다.

$$p(x_i | x_1, \ldots x_{i-1}) = p(x_i | \pi_i) \quad \ldots\ldots (2.3)$$

각 노드의 확률을 데이터로부터 학습하게 되는데 식 (2.4)와 같이 베이지안 네트워크를 수식으로 표현할 수 있다.

$$p(x) = \prod_{i=1}^{n} p(x_i | \pi_i) \quad \ldots\ldots (2.4)$$

변수 A, B, C, D, E, F가 있고 각 변수 들이 yes와 no 값을 갖는다고 할 때, 베이지안 네트워크는 그림 2-1과 같이 각 변수들에 대한 의존성을 그래프로 구성 할 수 있다. 그리고 각 변수 별로 표 2-1과 같은 조건 확률도 표현 할 수 있다.

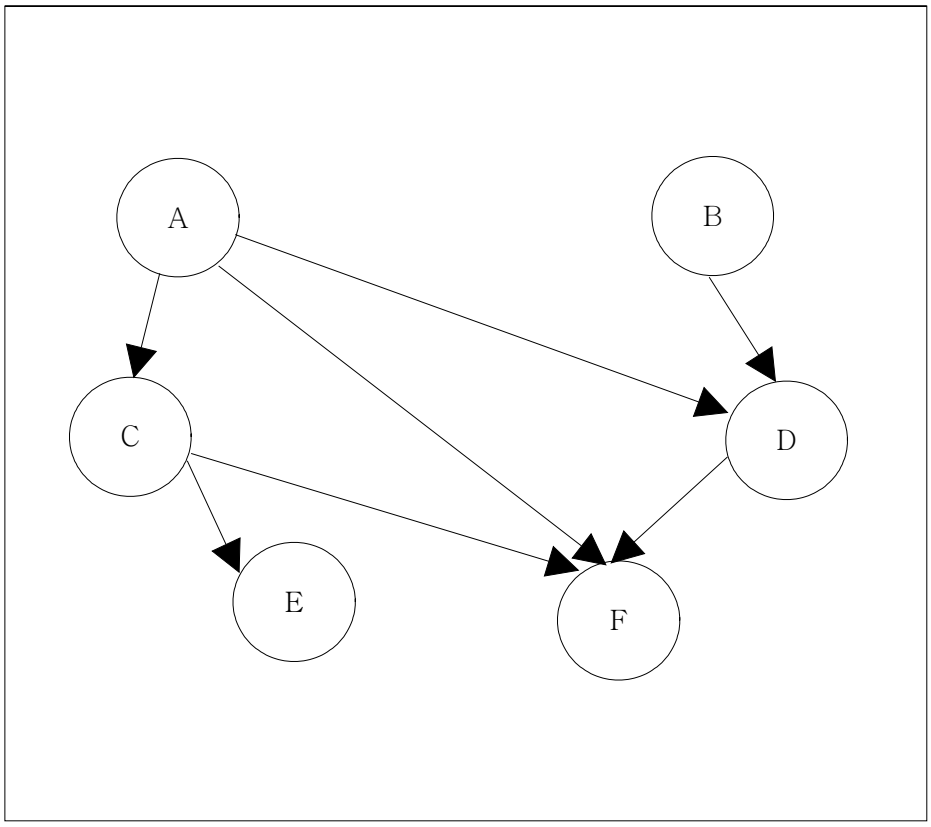


그림. 2-1 베이지안 네트워크 그래프

Fig. 2-1 Bayesian network graph

표 2-1은 6개의 변수 중 D가 각각 yes와 no 값을 가질 하나의 조건 확률을 나타내고 있다.

표. 2-1 조건 확률 표

Table. 2-1 Conditonal Probability Table

	A(y), B(y)	A(y),~B(n)	~A(n), B(y)	~A(n), ~B(n)
D(y)	0.4	0.1	0.8	0.2
~D(y)	0.6	0.9	0.2	0.8

그림 2-1에서 변수 D는 변수 A와 B에 대해서만 종속성을 갖는다. 따라서 표 2-1의 조건 확률 표의 각 열은 이 두 변수 A와 B에 배정 가능한 값들을 나타내고, 각 행은 변수 D가 yes 또는 no를 가질 조건 확률을 나타낸다. 변수 A가 no, B가 yes일 때 변수 D가 yes일 확률은 0.8이다. 반면에 D가 no일 확률은 0.2이다. 즉, 두 조건 확률은 각각 P(B=yes|A=no, B=yes)=0.8, P(D=no|A=no, B=yes)=0.2이다. 베이지안 네트워크는 조건 확률 계산 시 식(2.5)와 같은 베이즈 이론을 이용한다. 베이즈 이론은 관측된 데이터 D로부터 가설 h가 옳을 확률 P(h|D)를, 가설 h에 기초한 데이터 D의 조건 확률 P(D|h)과 가설 h의 사전확률 P(h)을 기초로 계산할 수 있는 방법을 정의한다[9][10]. 주어진 학습 데이터 집합을 D라고 할 때 가설 h에 대한 베이즈 이론에 의한 사후 확률은 식(2.5)와 같다.

P(h|D)는 h의 사후 확률이고, 이를 구하기 위해서는 P(D|h), P(h), P(D)를 먼저 구해야 한다.

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \quad \text{.....(2.5)}$$

또한 최대 사후 확률은 식 (2.6)과 같다.

$$h_{MAP} \equiv \underset{h \in H}{\arg\max}\, P(D|h)P(h) \quad \text{.....(2.6)}$$

$$h_{ML} \equiv \underset{h \in H}{\arg\max}\, P(D|h) \quad \text{.....(2.7)}$$

사전 가설의 확률이 없을 때 모두 같다고 가정하고 P(h) 값을 생략 한 후 최대 유사 확률은 식 (2.7)과 같이 표현한다.

2.2 눈동자 움직임을 이용한 시선 식별 법

사용자의 눈동자 움직임을 추적하여 시선을 식별하는 연구방법은 일반 조명 환경에서 획득한 영상을 이용하는 연구와 적외선과 같은 특수 조명 환경에서 획득한 영상을 이용하는 방법 등으로 나눌 수 있다. 또한 일반 조명하에서 영상처리방법[11]으로 눈동자의 움직임을 추적하는 연구들은 눈동자의 홍채(iris)를 추적하는 연구와 홍채내의 동공(pupil)을 추적하는 방법 등으로 나눌 수 있다[12].

그림 2-2와 같이 눈 안의 흰색부분 공막과 회색부분 홍채 사이의 경계를 영상처리 방법으로 추적하거나[13], 홍채 자체의 움직임을 입력으로 한 신경망을 이용하여 시선 위치를 파악하는 연구들이 있다[14].

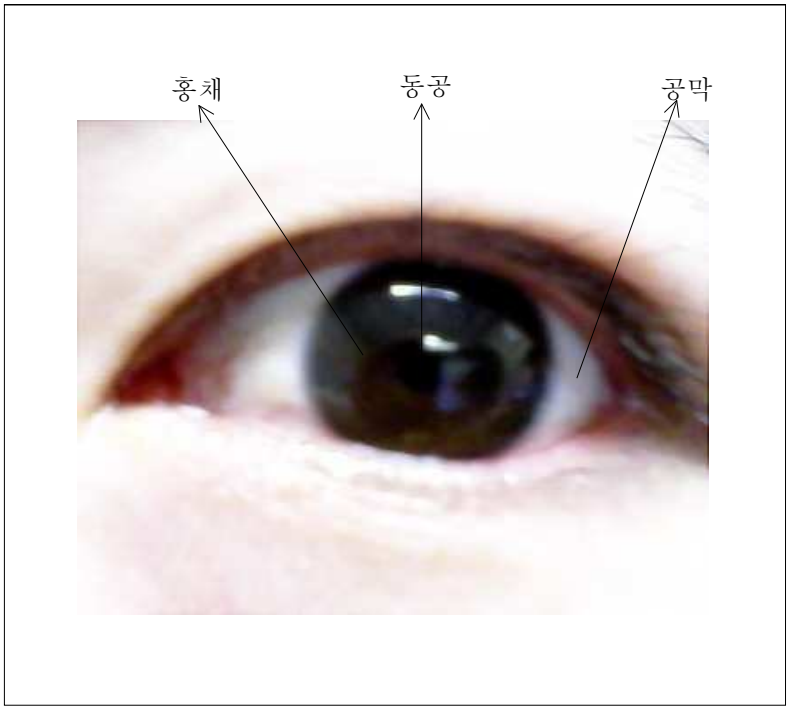


그림. 2-2 눈동자내의 홍채, 동공 및 공막 영역

Fig. 2-2 Iris, pupil and sclera in eye

이 방법은 공막과 홍채 경계의 위, 아래 부분은 눈꺼풀에 의해 덮여 있는 경우가 많아, 상하 방향의 눈동자 움직임은 정확하게 추적하기 어려운 단점이 있다[15]. 또한, 경계가 번짐 현상 등으로 명확하게 나타나지 않거나, 잡음의 영향을 잘못 형성되는 경우에 제대로 추출되지 않는 경우가 빈번히 발생한다. 이외에도 가변 템플릿 정합 방법을 이용하여 홍채의 중심위치를 찾음으로써 시선 위치를 파악하는 연구들이 있다[16]. 이 방법은 홍채와 공막사이의 경계가 명확하지 않거나 잡음이 어느 정도 섞여도 영향을 덜 받는 장점은 있으나, 템플릿의 초기 위치에 따라 수렴하는 속도가 많은 차이를 보이는 단점이 있다.

홍채내의 검은 부분인 동공을 영상처리 알고리즘으로 추적하는 방법은 동공이 눈꺼풀에 의해 덮여 있을 확률이 낮으므로 수평 방향 및 수직방향의 추적이 가능하다[17]. 앞에서 기술한 홍채 및 동공 영역을 추적하는 방법들은 일반 조명에서 영상 처리 방법에 의해 눈동자의 움직임을 추적하므로 특수 조명이 필요 없고 손쉽게 구현할 수 있다는 장점이 있지만 동공과 홍채사이의 계조도 차이가 아주 작기 때문에 정확한 추출이 어렵고 정확도가 낮은 단점이 있다[18]. 따라서 적외선과 같은 특수 조명을 이용하여 동공의 움직임을 추적함으로 추출이 어려운 단점을 극복한다.

2.2.1 적외선 기반 눈 검출 및 추적

지금까지 얼굴 및 눈동자 움직임을 함께 고려하여 시선위치를 파악하는 연구는 거의 이루어지지 않고 있다. Ohmura와 Ballard 등의 연구에서는 초기에 얼굴 특징점을 3차원 거리 정보를 알고 있어야 하며[19], 시선 위치를 파악하기 위해서는 많은 처리시간이 소요되는 단점이 있었다[20]. Gee와 Heinzman의 연구에는 얼굴 좌표계에서의 시선 벡터의 방향을 계산하였을 뿐[24], 이로부터 모니터 상에 사용자 시선위치 등을 구하지 않았다[21]. Rikert의 연구는 학습 및 테스트 환경에서 얼굴 및 모니터 스크린사이의 거리가 변하지 않아야 한다고 가정했으며, 이 경우 실제 사용시 많은 불편함을 제공하게 된다[22]. 기타 다른 연구들에서는 사용자로 하여금 구분점이 부착된 안경을 착용하게 함으로써 얼굴 특징 점을 추적하는 연구를 수행 하였다[23]. 이처럼 별도의 안경을 착용해야 하므로 사용자에게 불편을 제공하는 결과가 되었다. 이와같은 기존의 대부분 연구는 영상 처리 방법에만 의존하여 시선 위치 추적 연구를 수행하였기 때문에, 처리 속도도 늦고 많은 사용 제약을 갖는 문제점이 있다.

이러한 문제점을 해결하기 위해, 사용자가 장비 내부의 카메라에 광축을 맞추어 접근할 때 자동으로 동공인식에 적합한 눈 영상을 획득하여 인식하도록 하는 것이다[24]. 즉, 적외선 동공 인식방법은 흑백 고정초점 CCD카메라, 적외선 조명, 반투명 거울을 이용하여 자신의 눈을 장비 내부의 카메라와 광축에 쉽게 맞출 수 있고, 반투명 거울과 카메라 사이의 적외선 필터로 인해 불필요한 가시광선이 영상으로 들어오는 것을 막을 수 있다. 적외선을 이용하여 동공인식에 적합한 눈 영상을 획득하기 위하여 사용자는 다음과 같은 절차를 따른다. 사용자는 장비에 접근하여 자신의 눈을 장비의 반투명거울에 표시된 표시 자에 일치시켜 광축을 맞춘다. 그 상태에서 장비가 있는 쪽으로 접근하면 시스템은 자동적으로 동공인식에 가장 적합한 눈 영상을 획득하게 된다.

2.2.2 밝은 동공 효과와 글린트 추적

적외선을 눈동자에 투사했을 경우에 생기는 각막에서의 반사광(glint)과 밝은 동공효과(bright pupil)를 이용하여 눈동자의 움직임 및 시선 방향을 추출한다[25]. 이 경우 일반 조명을 사용 했을 때 보다 정확하게 눈동자를 찾을 수 있지만, 적외선과 같은 특수 조명이 필요하다는 단점이 있다.

앞에서 기술한 홍채, 동공이나 각막과 망막의 반사광을 추적하는 연구에서는 얼굴이 고정된 상태에서 단지 눈동자의 움직임만을 추적하거나 혹은 눈동자의 움직임으로부터 모니터 상의 시선 위치를 파악한다. 그러므로 얼굴의 움직임이 있는 경우에는 시선 식별의 어려움이 있고 각각의 사용자에 따른 교정 작업이 필요하다는 단점이 있다.

2.3 얼굴 움직임을 이용한 시선 식별 법

정적 혹은 동적인 배경에서 움직이는 물체를 분리하거나 검출하기 위해 동영상을 분석하는 방법으로는 차 영상의 분석에 기반 한 방법, 분류기법을 사용한 유동장의 형태를 계산하고 계산된 유동장에서 물체상의 가시점의 구조나 깊이를 측정하는 것 등이 있다. 이산적인 방법은 윤곽선, 에지, 모퉁이선, 궤적과 같은 특징들의 프레임의 순서에 따라서 추적되고, 이러한 특징들의 구조를 계산한다.

Li는 얼굴의 3차원 모델을 모델 기반의 압축 연구에 사용하였다[26]. 이러한 모델들은 목격자에 의한 범인 얼굴의 재구성이나 부분적 정보만을 사용한 얼굴 재구성, 컴퓨터에 의한 자동 얼굴 변형 시스템 같은 응용분야에 매우 유용하다.

Yang은 동영상을 사용한 얼굴 표정 인식 방법에 대해서 제안하였다[27]. 이 방법은 사람의 표정을 화남, 불쾌함, 슬픔, 놀라움, 공포의 여섯 가지로 분리하고, 동영상에서 시선을 식별하기 위해 얼굴의 특징을 분석한 후, 특징 점들의 위치 정보를 바탕으로 얼굴의 표정을 인식하였다. 30명의 개인에 대한 120개의 동영상에 대한 실험에서 이 방법은 매우 성공적으로 수행되었다.

얼굴의 삼차원 움직임 양을 구하는 연구는 그림 2-3과 같이 이차원 카메라 영상에 투영된 얼굴 특징 점의 움직임으로부터 얼굴의 삼차원 회전량 및 이동 량을 추정하는 방법들을 사용한다. 얼굴의 삼차원 움직임 양 추정을 위해서 Azarbayejan는 확장 Kalman필터를 Fukuhara는 신경망과 모델기반 영상 압축에서 필요로 하는 얼굴 회전량 및 이동 량을 신경망을 이용하여 추출하는 방법을 사용하였다[28].

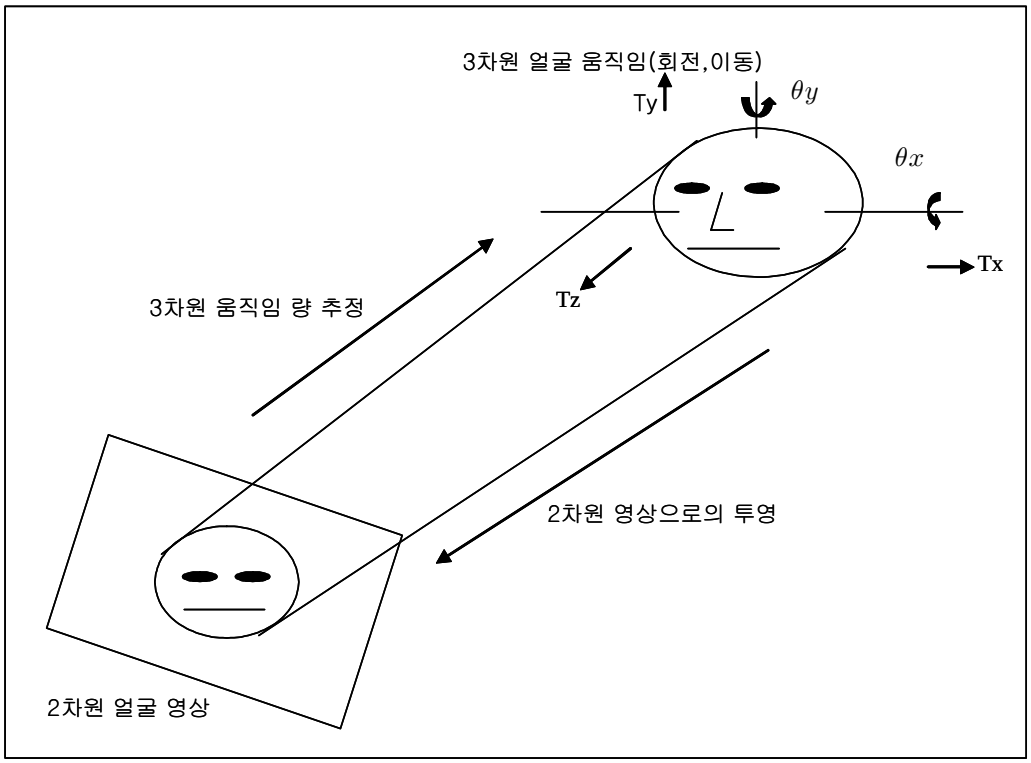


그림. 2-3얼굴의 3차원 움직임량 추정

Fig. 2-3 3D Faical motion value estimation

2.4 얼굴 특징 추적

얼굴의 특징을 추출하기 위한 방법으로 여러 가지 방법이 연구되고 있다.

Harmon은 옆 얼굴 사진에서 9개의 기점을 이용하여 기점 간 기울기, 턱이나 이마 곡면의 반경, 코의 돌출 정도, 얼굴의 면적 등 모두 11개의 특징 값을 이용하여 얼굴 인식을 시도하였다[29]. Wong은 정면 영상을 이용하여 얼굴 윤곽선 및 특징점간 거리 등을 이용하여 거리 값의 변화 정도에 따라 가중치를 적용하는 방법으로 인식을 시도하였다[30].

Sirovich 와 Kirby는 수염이 없는 얼굴 영상에 대해 KL 변환을 적용하여 얼굴을 표현하는 방법을 소개하였다[31]. 이 방법은 모든 영상이 KL변환에 의한 고유벡터 값의 가중치 조합을 사용하여 재구성 될 수 있으며, 영상의 재구성에 사용하는 고유벡터의 수를 증가시키면 입력 영상에서 보다 가까운 영상을 구성할 수가 있다는 점에 기반을 두고 있다. Yuille은 가변 형판을 사용하여 얼굴의 특징을 산출하였다[32]. 여기서 사용되는 형판은 해당 특징에 부합되기 위해서 변형, 회전, 이동할 수 있다.

Hallinan도 영상에서 눈을 검출하기 위해 템플릿 매칭을 사용하였다[33]. 이 방법에서 눈의 템플릿은 눈동자 영역과 흰자위 영역으로 구성되며, 두 영역은 단일한 명도를 갖는다. 이 방법은 107개의 눈 데이터, 그리고 눈과 비슷한 영상들을 사용한 실험에서 80% 정도의 성공률을 보였다[34]. Manjunath은 Gabor 웨이블릿 분해와 지역적 크기 변환의 상호작용을 사용하여 곡률이 최대인 점에서의 얼굴의 특징을 추출하는 방법을 제시하였다[35]. 이러한 특징들은 데이터베이스에 저장된 다음 그래프매칭 기법이 적용되었는데, 이 방법에서는 2차원의 Gobor 함수와 Fourier 변환도 사용되었다.

또 다른 연구접근 방법으로 신경회로망을 이용하여 특징 점을 찾은 방법이 있는데 British Telecom연구소와 Vincent는 다층퍼셉트론을 이용하여 특징 점을 찾는 방법을 사용했다[36].

본 논문에서는 얼굴의 진위 여부를 판단하기 위해서 얼굴의 특징 요소인 양쪽 눈, 양쪽 콧구멍 및 입의 양끝을 사용하여, 구조적 정보와 비교적 처리 시간이 빠른 수평, 수직 히스토그램 분석법을 이용하여 얼굴의 요소를 검출한다.

눈의 위치는 그림 2-4와 같이 전체 얼굴크기의 1/2정도 높이에 두며, 코끝의 위치는 눈 위치의 1/2정도 높이로 하고, 입은 코의 위치로부터 1/2정도 되는 높이에 위치한다.

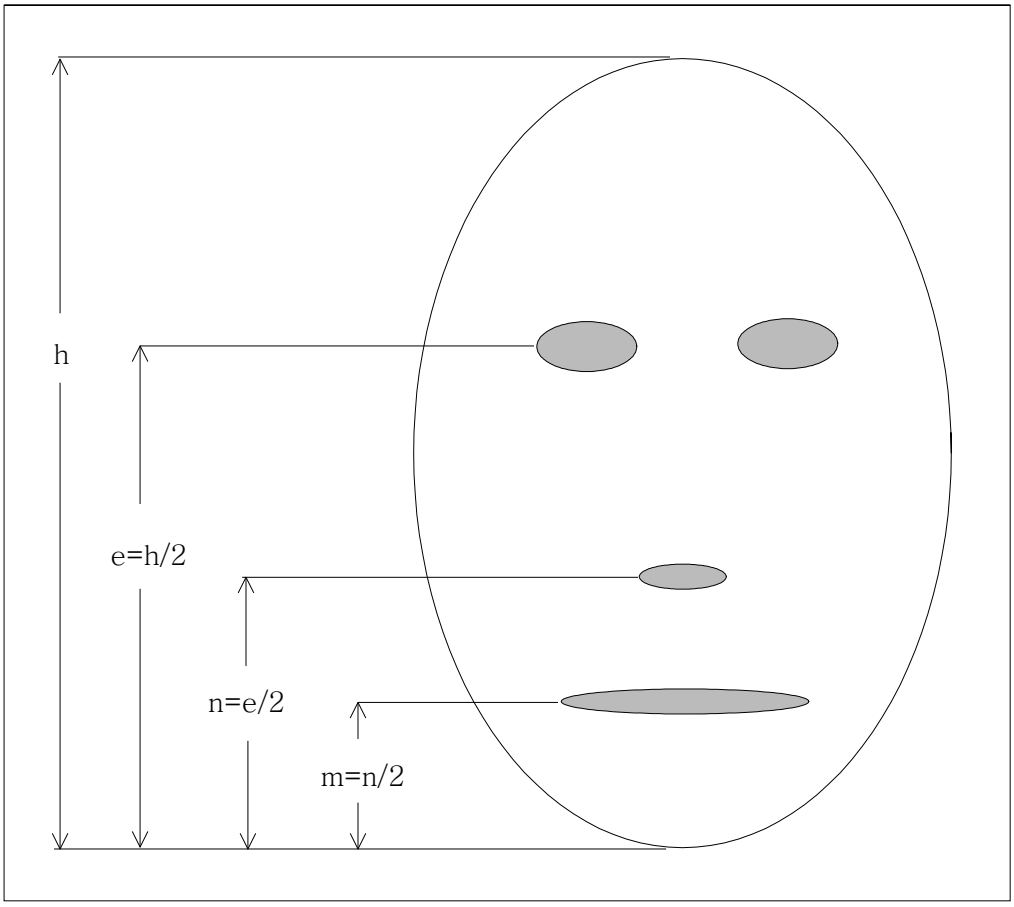
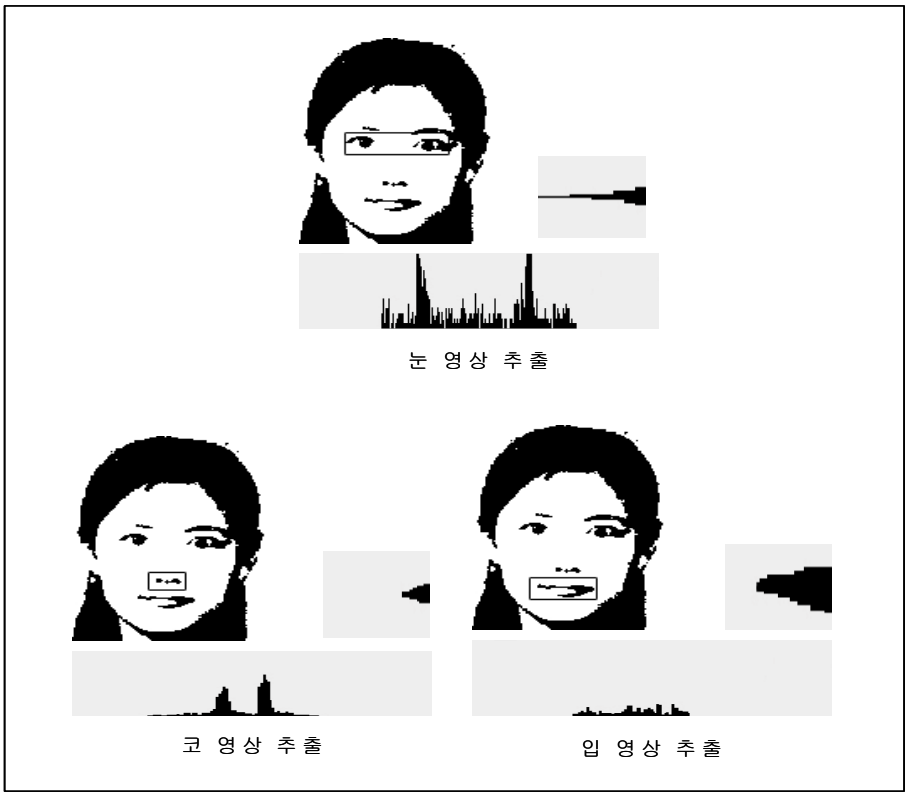


그림. 2-4 얼굴의 구조적 특징

Fig. 2-4 Structural feature of Face

이러한 구조적 정보를 바탕으로 얼굴요소의 후보 위치를 선정하고 그 위치 내에서 얼굴 요소를 추출한다. 얼굴 영역에 대한 이진영상을 구한 다음 얼굴의 구조적 정보를 바탕으로 눈이 존재한 것으로 예측되는 영역에서 수직, 수평 투영 히스토그램을 구한다. 눈이 존재할 경우는 그림 2-5와 같이 수평축과 수직축에서 히스토그램의 피크 위치가 나타나고 이 축이 만나는 두 점이 양 눈의 위치가 된다. 눈 위치 추출 시 안경 사용자의 경우, 안경 렌즈의 반사에 의해 눈동자 흑화소가 많이 소실될 수 있으므로, 눈 영역의 이진화시 p-tile방법을 사용하여 흑화소의 개수를 일정하게 유지함으로써 흑 화소 소실 문제를 해결한다.



2-5 얼굴 특징 요소 추출

Fig. 2-5 Face feature point extraction

콧구멍 역시 눈동자와 같은 방법으로 추출하였고, 두 콧구멍의 근접도가 큰 관계로 양 콧구멍을 포함한 영역에 대한 수직 히스토그램을 통해 두 개의 피크 위치를 추출함으로써 양 콧구멍의 수직 위치를 파악한다.

또한, 얼굴의 구조적 정보를 이용하여 입의 위치에 대한 존재 가능 범위를 설정한 후 이 영역에 대해 수직, 수평 히스토그램을 구한다. 입의 특성상 수직 방향 히스토그램은 일정 부분에만 존재하게 된다. 임계치 이상으로 갑작스럽게 변하는 부분을 입의 양 끝점 위치가 된다. 이렇게 얻어진 얼굴 요소의 존재 유무에 따라서 얼굴의 진위여부를 판단하며, 정면 얼굴을 대상으로 하는 얼굴 검출의 결과는 두 눈 중심점 위치의 정확도로 평가된다. 사람의 얼굴은 모두 유사한 구조를 유지하기 때문에 두 눈의 위치 정보만 주어진다면 영상에서 얼굴인식에 필요한 영역을 추출할 수 있다. 따라서, 얼굴 특징요소 추출을 통해서 얻어진 눈의 위치를 중심으로 일정크기의 영상을 획득하게 된다.

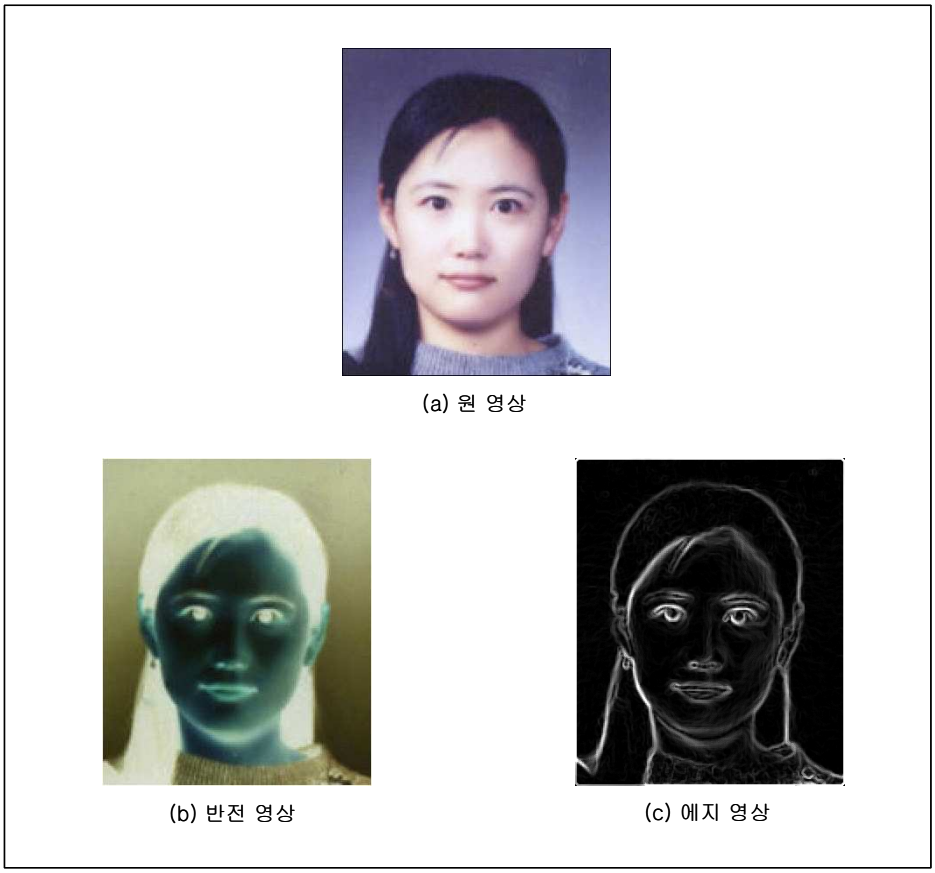
또한 얼굴 영역에서 머리카락 영역까지 얼굴 영역으로 검출되는 경우에는 눈이나 눈썹의 검출에 심각한 영향을 줄 수 있다. 따라서 머리카락 영역은 얼굴 영역에 비해 이웃 화소 간 밝기 성분이 크게 변하는 특징을 갖고 있고 식 (2.8)과 같은 분산 $V(x,y)$으로 표현 할 수 있다. 이 값이 정해진 문턱치 이상이면 머리카락 영역으로 표시한다.

$$V(x,y) = \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1} | Y(x,y) - Y(x+i,y+i) | \quad (2.8)$$

$$f_{RF}(x,y) = f_F(x,y) \cap \overline{f_H(x,y)} \quad (2.9)$$

머리카락 영역을 나타내는 영상 $f_H(x,y)$는 머리카락 영역은 255이고, 나머지는 0으로 표현된다. 기존 얼굴 방법으로 얻어진 $f_F(x,y)$과 머리카락 영상 $f_H(x,y)$을 식 (2.9)의 논리적 관계식으로 머리카락의 영향을 받지 않는 실질적인 얼굴 영역 영상 $f_{RF}(x,y)$을 얻을 수 있다.

그림 2-6는 임의의 샘플영상에서 얼굴 영역 검출 결과를 보여주고 있다. (b)는 단순히 색상 정보만으로 검출된 얼굴 영역으로 머리카락 영역을 포함하고 있고, (c)는 식 (2.9)를 이용해 검출된 얼굴 영역으로 많은 영역 중 가장 큰 영역만을 최종 얼굴 영역으로 선택한다.



(a) 원영상 (b) 반전 영상 (c) 에지 영상

그림. 2-6 얼굴 영상에서 얼굴 영역 검출

(a) input image (b) solarization image (c) edge image

Fig. 2-6 Face region detection

2.5 얼굴 동작 인식을 위한 인공신경망

인공신경망은 학습 방법에 따라 지도 학습법과 자율학습법의 2가지로 구분할 수 있다. 본 연구는 지도학습법의 GRNN 알고리즘으로 얼굴인식을 연구한다.

GRNN(Generalized Regression Neural Network)은 Specht에 의해 처음 사용된 것으로 관측된 데이터의 확률 밀도 함수에 기본을 두고 있다[37]. GRNN의 원리는 PNN(Probability Neural Network)의 기본 원리와 같고, PNN은 분류에 주로 사용되지만 GRNN은 매핑 기능에 이용된다. GRNN의 구성은 4개의 층으로 구성된다. 입력층(input layer)은 입력 패턴을 각각의 숨김층(hidden layer)의 뉴런들에 배분하는 역할을 하며 두 번째 층으로 모두 연결된다.

신경망 중 가장 널리 쓰이는 것은 BP(Back Propagation)이므로 축색돌기(axon)에 해당하는 뉴런(neuron)으로 연결된다. 뉴런에 입력된 입력 신호들은 각각의 입력 신호 값과 관련된 가중치들에 의해 곱해진 후 합해지게 되고, 이 합은 비선형적인 활성화 함수에 넣어 출력 값으로 나타나게 된다. 이때 BP에서는 식(2.10)과 같은 시그모이드(sigmoid)함수를 사용한다.

$$f(x) = \frac{1}{(1+e^{-x})} \quad \text{.....(2.10)}$$

식 (2.10)에서 계산된 출력 값은 뉴런의 위치에 따라서 다음 층의 입력 데이터나 최종 출력 값으로 사용된다. 인공신경망은 학습 데이터들을 사용하여 학습을 시키는데 인공신경망에서 계산된 값과 실제로 요구되는 측정값과의 오차의 자승으로부터 각 연결 가중치를 변화시킨다. 이것은 오차함수(error function)의 국부 최소 값에 수렴할 수 있고 값은 반복 계산이 필요하다. 그러나 BP는 학습 시간이 많이 소요되는 단점이 있다[38]. 이러한 단점을 보완하기 위한 수단으로 다음과 같은 두 가지 방법을 들 수 있다. 첫째, 학습 시간을 단축시키는 방법으로 운동량을 더하는 방법과 학습률을 변화시키는 방법이 있다. 둘째, 국부 최소값에 수렴하지 않도록

하기 위한 수치적인 적합화 방법으로 직교기울기법(orthogonal gradient method)과 Levenberg-Marquart방법을 사용하기도 하지만 여전히 많은 문제점이 존재하게 된다. 네트워크 진행 방향이 어떻게 될 것인가에 대한 이론적인 예측이 매우 힘들고, 초기 가중치와 숨김 층에 있는 뉴런의 개수에 따라 많은 수의 가중치를 구해야할 경우에는 적절한 결과를 얻기 힘들다.

숨김 층은 학습 데이터에 있는 각각의 패턴들에 대해 하나의 뉴런을 가지고 있으며 이때의 가중치는 학습 데이터에 있는 패턴 요소들이 된다. 각각 숨김 층 뉴런들이 가지고 있는 학습 데이터로부터 입력 층 원소들을 빼어 제곱을 하거나 절대값으로 나타낸다. j번째 숨김 층의 입력 함수는 식 (2.11)와 같다.

$$I_j = \sum_{i=1}^{n} |W_{ij} - X_{ij}| \quad \text{또는} \quad I_j = \sum_{i=1}^{n} (W_{ij} - X_{ij})^2 \qquad (2.11)$$

여기에서 x_i는 입력신호이고 w_{ij}는 i번째 입력 층과 j번째 숨김 층 뉴런 사이의 가중치이며 n은 입력 층의 뉴런 수를 나타낸다. 이렇게 계산된 입력 값은 식 (2.12)과 같은 지수함수의 형태를 가지는 비선형 활성함수로 넘겨지게 된다.

$$f(I_i) = \exp\left(\frac{-I_j}{2\sigma^2}\right) \qquad (2.12)$$

활성 함수에서 계산된 숨김 층의 출력 값은 합산 층(summation layer)으로 전달되며, 합산 층에서는 식(2.13)와 같은 적분을 실행하여 결과 값을 계산하여 출력 층(output layer)으로 보낸다.

$$Y(X) = \frac{\sum_{i=1}^{n} Y' f(I_i)}{\sum_{i=1}^{n} f(I_i)} \qquad (2.13)$$

여기서 Y'는 각 숨김 층 뉴런이 가지는 출력 값으로써 학습 데이터의 결과 값에 해당한다. 위의 과정들을 정리하면 다음과 같다.

GRNN은 각각의 학습 단계에서 네트워크에 입력하여 새로운 입력 값에 대한 반응을 계산하는데 사용한다. 그리고 가우시안 커널(Gaussian Kernel)함수들을 각 학습 단계에 위치시킨다. 입력 값으로 계산된 결과 값은 학습 단계의 결과 값 평균에서 가중치를 두어 계산되며, 가중치는 계산 값과 학습 단계와 관련이 있다. 즉, 입력 값에서 가까운 학습 단계의 결과 값은 계산에 있어서 영향을 미치게 된다.

역전파에 비해 GRNN이 가지는 큰 장점은 학습 데이터가 비교적 적게 요구된다는 점이다. 역전파와 비슷한 정확도를 얻기 위해 약 1%의 학습 데이터가 필요하며 [39], 학습 데이터가 많은 경우에는 비슷한 데이터들을 군집화 시켜 사용할 수 있다.

III. 개선된 Kalman 필터 알고리즘 제안

3.1 Kalman 필터 이론

Rudolf Emil Kalman은 “상태변수의 개념을 Wiener filtering problem에 적용할 수 없을까?”라는 의문점에서 Kalman 필터의 연구를 시작하였다[40]. Rudolf Emil Kalman은 1960년 이산 데이터 선형 필터링 문제에 대한 재귀적인 해에 관한 논문을 발표하였다. 이산 Kalman 필터는 이론적으로 선형 2차 가우시안 문제에 대한 하나의 추정기이며, 이는 가우시간 백색잡음에 의해 교란된 선형 동적 시스템의 순간 상태를 추정하는 것이다.

동적 시스템을 제어하기 위해서는 무엇이 행해지고 있는가를 먼저 인식해야 한다. 제어하고자 하는 모든 변수는 측정 가능하지 않으며, 또 모든 변수를 측정하는 것이 바람직한 것은 아니다. 이때 Kalman 필터는 간접적인 방법으로 측정 불가능한 정보를 이해할 수 있도록 해 준다. 주어진 기본 추정 값에서 모델 변수들이 추정되고 이는 새로운 추정 값으로 조정되게 하며, 또 각각의 갱신 시에 그 에러에 대한 추정을 제공한다. 시간에 대한 측정값들의 집합으로 시스템의 상태를 추정할 때 Kalman필터는 상태변수들의 간접적인 측정값과 상태변수 및 간접적인 측정치 둘 모두의 공분산 정보를 사용하여 시스템 상태의 이전 추정 값을 갱신하거나, 추정하는데 까지 연구한다. 그림 3-1은 Kalman필터 loop를 보여주는 것으로, 반복적인 과정을 거치면서 추정 값을 갱신하는 것을 나타냈다.

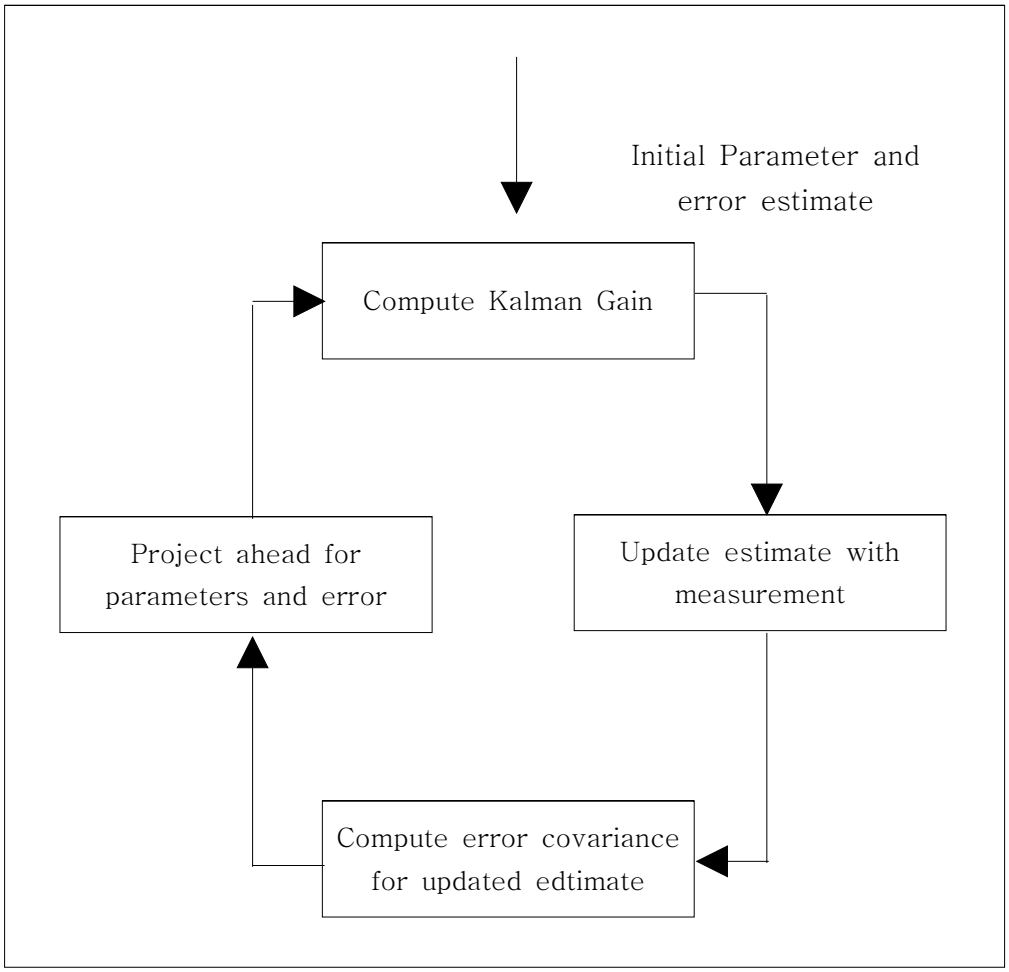


그림. 3-1 Kalman 필터 루프
Fig. 3-1 Kalman filter loop

Kalman필터는 초기변수들, 변수와 관련된 에러의 추정 값, 측정 에러의 분산을 포함하는 초기 값을 가지며, 이러한 값들은 게인 행렬을 구하는데 사용된다. 변수 추정 값과 측정된 데이터 사이의 에러는, 변수 추정 값과 추정에러를 갱신하기 위해 게인행렬(gain matrix)에 의해 곱해지고, 갱신된 에러와 변수는 모델의 입력으로 사용된다. 모델 변수들이 나쁜 추정 값들을 가질 때, 게인 행렬은 측정데이터가 상태변수를 추정 시에 매우 영향을 받기 쉽다. 그리고, 반복에 따라 변수의 정확도에 대한 신뢰성이 증가하면 게인 행렬 값은 감소하고, 변수와 그 관련 에러의 갱신에서 측정 데이터의 영향을 작아지게 한다.

3.1.1 Kalman 필터 변수 및 방정식

$q(k)$: 시간에서 t_k 현재의 변수를 포함하는 벡터.

$\bar{q}(k)$: $q(k)$와 $Y(k)$사이의 에러를 갱신하기 전, 시간 t_k에서 변수의 현재 추정치를 포함하는 벡터.

$\bar{q}(k+1)$: 다음 시간 샘플에서의 변수 추정치 벡터.

$p(k)$: 현재 변수들의 에러 공분산 행렬.

$M(k)$: 변수 추정치들의 에러 공분산 행렬.

$M(k+1)$: 다음 순간의 변수 추정치의 에러 공분산 행렬.

$Y(k)$: 최근의 위치 측정값을 포함하는 벡터

C : 시간 t_k에서 상태벡터와 측정사이의 노이즈 없는 연결을 제공하는 행렬.

R_v : 측정에러 공분산 값을 나타내는 행렬. 알고 있거나 추정되어 진 값이어야 한다.

R_w : 시스템의 프로세스 노이즈 공분산 행렬.

A : $x(k)$에서 $x(k+1)$로의 관계를 나타내는 모델을 포함하는 행렬.

$G(k)$: $\bar{q}(k)$와 $Y(k)$사이의 에러의 영향의 정도에 관계된 게인 행렬.

① Computer Kalman gain

$$G(k) = M(k)C^t[CM(k)C^T + /R_v] - 1 \quad \text{.....(3.1)}$$

측정에서 얻는 상태변수 벡터에서의 추정된 에러 값을 사용하여 게인 행렬 $G(k)$가 계산된다[41]. 이들 매개변수의 추정에러는 측정에러를 더한 값으로 나누어진다. 이것은 $M(k)$가 측정에러에 대하여 더 작은 연관성을 가질 때 게인 행렬 $G(k)$값의 영향을 줄이는 효과가 있다.

$G(k)$는 추정변수와 측정변수사이의 에러 비를 결정하며 상태벡터를 갱신할 때 사

용되며, $M(k)$에러가 감소할 때 측정 데이터의 영향 또한 감소하는 것을 보여 준다.

② Update estimate with measurement

$$q(k) = \bar{q}(k) + G(k)[Y(k) - c\bar{q}(k)] \quad (3.2)$$

매개변수의 추정값과 측정값의 차이는 추정에러로 생각할 수 있으며 잘 못된 측정, 잘못된 추정, 혹은 두 가지 모두의 결과로 발생된다. 이는 관련 측정 데이터에서 추정 상태매개변수 데이터를 빼는 것으로 유도된다.

갱신된 상태 매개변수 $q(k)$벡터를 만들기 위해 추정에러의 비율이 매개변수 추정값 $\bar{q}(k)$에 더해진다. 비율은 게인 행렬 $G(k)$의 값에 의해 정의된다. 이는 기존의 측정 에러와 마찬가지로 상태 매개변수에서의 현재 신뢰도에 의해 정의된다.

③ Computer updated erroe convariance

$$P(K) = M(k) - G(k)CM(k) \quad (3.3)$$

비게인의 비율을 나타내는 단위 행렬에서 측정으로 얻은 요소와 관련된 게인 행렬의 몫을 뺀다. 이는 갱신된 에러분산을 계산하기 위한 추정에러 분산으로 곱해진다. 즉, $G(k)$가 감소할 때 $p(k)$를 갱신하는데 사용되는 추정에러의 양이 비례하여 증가한다. 극단적으로 $G(k)$ 행렬이 단위 행렬이라 하면, 전체 예측에러를 식 (3.2)에 의해 상태 매개변수에 더하면, 측정된 데이터와 정확히 정렬되게 한다. 이 경우 $p(k)$는 식 (3.3)에서 계산되며 식 (3.2)에서의 측정 정보에 맞는 상태변수 벡터에 대하여 명확성을 보여준다. 상태 매개변수 벡터에서 다른 변수들에 대해서도 불확실성은 그대로 존재한다.

④ Project ahead for parameter and error

$$\bar{q}(k+1) = Aq(k) + Bu(k) \quad (3.4)$$

$$M(k+1) = AP(k)A^T + BR_wB^T \quad (3.5)$$

다음 순간의 상태 매개변수 값과 에러의 예측은 식 (3.4)와 (3.5)에 의해 이루어진다. A모델은 현재의 매개변수와 에러에 대한 근사함수를 적용하여 유도되는 새로운 추정치를 나타낸다. R_w행렬은 모델링 에러를 나타내는데 사용된다.

3.2 Kalman 필터 알고리즘

JI가 제안한 Kalman 필터에 의한 동공 추적 알고리즘은 순환적 필터 방식으로 알고리즘이 단순하다[42]. 또한, 시간영역에서 직접적인 처리가 가능하므로 디지털 컴퓨터에서 사용하기 적합하다. Kalman 필터는 상태변수 측정 시 포함된 노이즈를 감소시키고 상태변수를 추정하는 역할을 하고, 미지의 플랜트 파라미터를 추정할 수 있다. 상태 방법 식은 식(3.6)과 같다.

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_k + w_{k-1} \\ y_k &= Cx_k + v_k \end{aligned} \quad \cdots\cdots (3.6)$$

식(3.6)에서 $x(k)$는 시스템 상태를, $y(k)$는 측정값을 나타낸다. $w(k)$와 $v(k)$는 프로세스와 측정 노이즈이다. A, B, C는 적당한 차원을 갖는 알고 있는 실수 행렬이므로 Kalman필터 알고리즘은 다음과 같다.

① 주어진 행렬 A와 시스템 상태로부터 초기 상태 값을 추정한다.
추정된 상태 값으로 부터 다음 순간의 에러 값을 예측한다.
② Kalman 게인을 계산한다.
③ 측정된 추정치를 갱신한다.
④ 에러 공분산을 갱신한다.
⑤ ① ~ ④ 반복한다.

그림 3-2는 물체추적시스템에서 Kalman 필터를 이용한 알고리즘을 나타낸 것으로 Kalman필터는 물체를 윈도우 중심에 두기 위해서 이전의 데이터로 다음 샘플링 시간에서의 위치를 예측하는 방법을 이용한다.

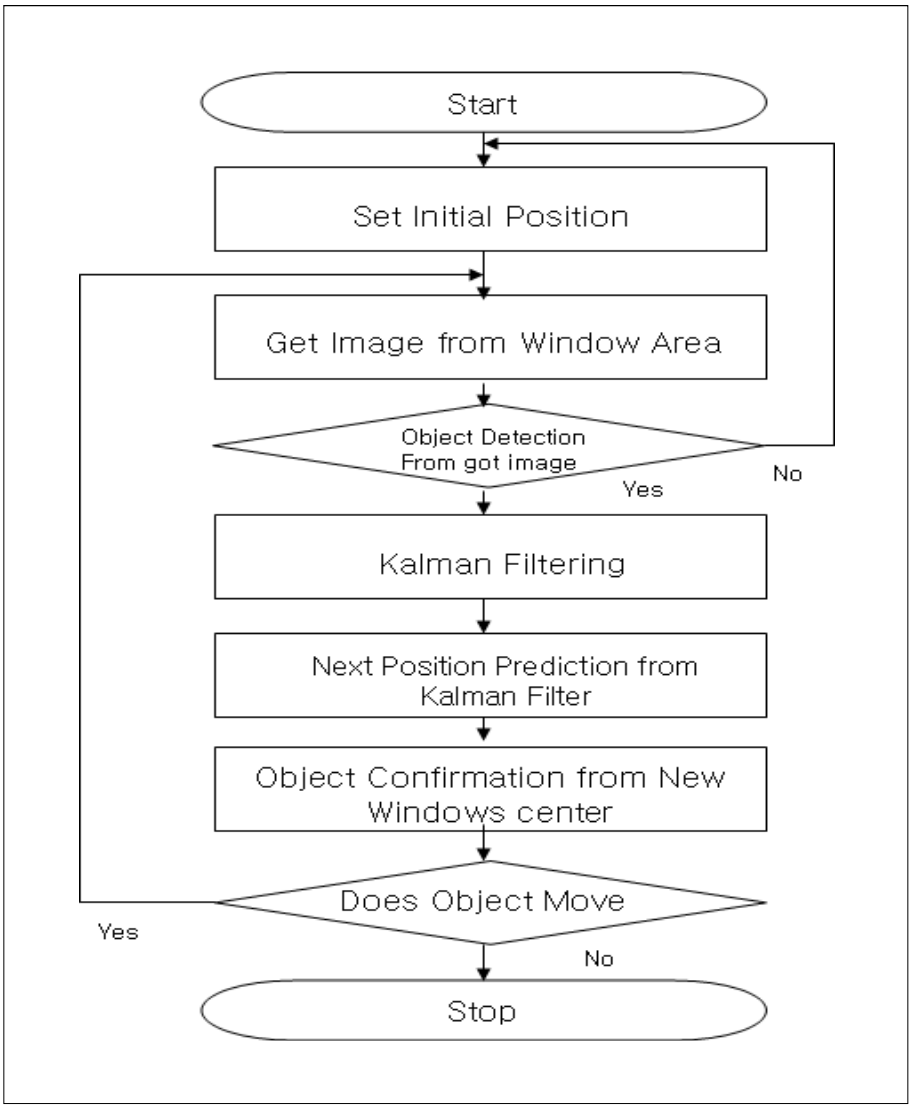


그림. 3-2 Kalman필터 알고리즘
Fig. 3-2 Kalman filter algorithm

3.3 Kalman 필터와 평균 이동 알고리즘을 이용한 동공 추적

Kalman 필터를 기반으로 한 밝은 동공을 이용한 눈 추적기를 이용하여 초기 프레임에서 동공의 위치를 확인하고 Kalman필터에 의하여 밝은 동공을 추적하였다.

각각의 프레임에서의 동공의 움직임은 동공의 위치한 변화속도로 특성화 된다. 시점 t에서 동공 중심의 화소의 위치를 $(c_{t,} r_t)$라 하고 (u_t, v_t)를 시점 t에서 c와 r 방향으로의 변화속도라 하면, 시점 t에서 상태벡터를 $x_t = (c_t, r_t, u_t, v_t)t$로 나타낼 수 있다. 그리고 시스템을 다음과 같이 모델링 할 수 있다.

$$X_{t+1} = \Phi X_t + W_t \quad (3.7)$$

여기서 W_t는 시스템의 불안전성을 나타낸다. 시점 t에서 동공의 위치를 $Z_t = (\hat{c_t}, \hat{r_t})$로 간주하면, Kalman 필터에 의한 측정모델을 구할 수 있다.

$$z_t = Hx_t + v_t \quad (3.8)$$

여기서 v_t는 불확정된 측정치이다. 현재 프레임 t의 위치는 시스템모델을 기반으로 밝은 동공효과가 존재할 만한 예상 위치를 구하고 인접한 화소에서 문턱치값을 이용하여 추정한다. 그리고 식 (3.7)의 상태 모델과 식(3.8)의 측정 모델을 통하여 상태 벡터 x_{t+1}와 공분산 행렬Σ_{t+1}의 초기치를 얻고 예측을 위한 시스템 모델과 측정 모델을 갱신하게 된다.

연속되는 두 프레임에서 한 사람의 동공은 상당한 위치적 변화를 갖지 않으므로, 이후 프레임에서 얼굴 위치를 예측하기 위해[42], 이전 프레임의 동공 위치를 기초로 Kalman필터링을 사용하여 추적이 보다 효율적으로 이루어질 수 있도록 하였다[43].

Kalman 필터로 동공 추적 시 눈이 감기거나 폐색에 의하여 밝은 동공이 사라질

경우 Kalman 필터에 의한 추적이 불가능해지는 단점이 있으므로, 평균 이동 알고리즘을 이용하여 눈 추적의 정확성을 높였다. 동공의 위치를 확인 후 Kalman 필터에 의하여 동공 추적이 시작된다. 그러나 눈이 감기거나 폐색에 의하여 밝은 동공 효과가 나타나지 않을 경우 평균이동 추적에 의한 동공 추적이 시작된다. 그리고 이후 밝은 동공이 다시 나타나면 밝은 동공을 이용한 추적으로 즉시 돌아간다. 제안된 동공추적 알고리즘을 그림 3-3에 나타냈다.

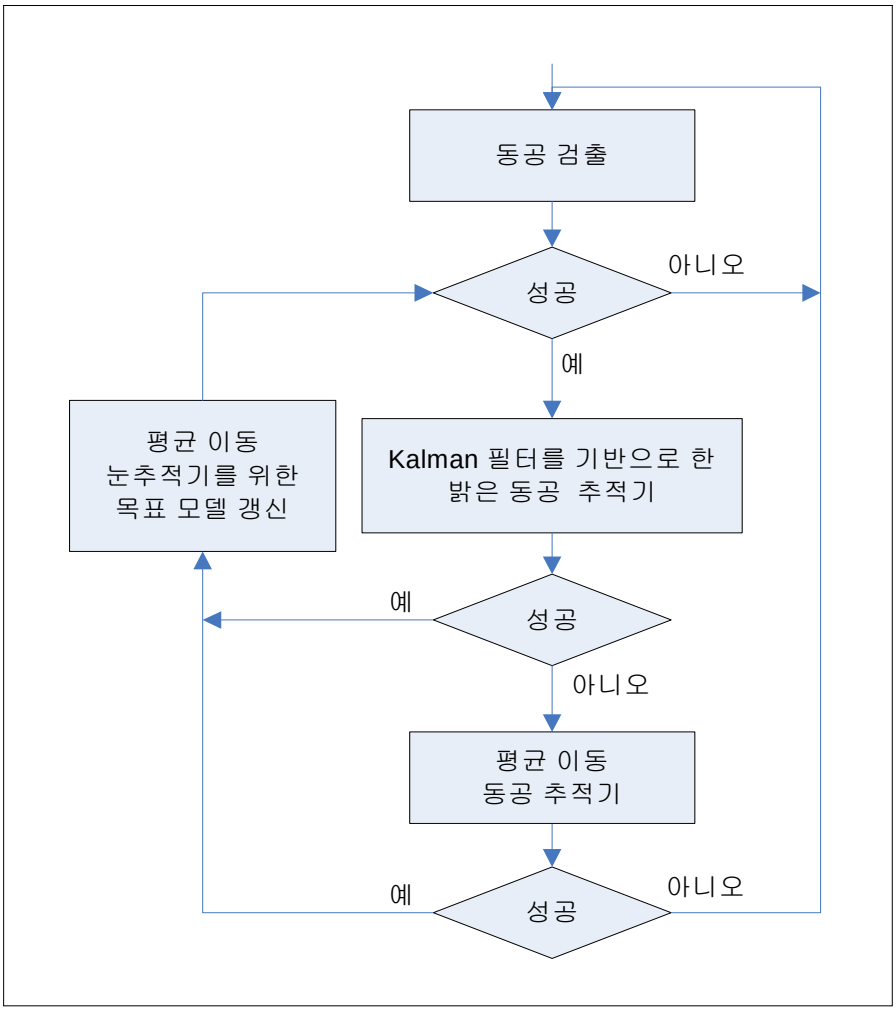


그림. 3-3 동공 추적 알고리즘

Fig. 3-3 Pupil tracking algorithm

평균 이동 추적 알고리즘은 바타라치야 계수로 두 개의 명암 분포의 유사성을 측정하여, 유사성에 따라 평균 이동을 반복하여 모델과 가장 유사한 후보 영역을 찾은 것이다. 견본 데이터로부터 유도된 바타차리야 계수는 표적 명암 q와 후보 명암 p의 분포에 영향을 미친다[44]. y지점에서 표적 밀도q를 위한 바타차리야 계수를 추정하고, 표적 후보 밀도 p(y)를 식(3.9)에 의하여 구하였다.

$$\hat{p}(y) \equiv p[\hat{p}(y), \hat{q}] - \sum_{u=1}^{m} \sqrt{\hat{p}_u \hat{q}_u} \quad \cdots\cdots (3.9)$$

여기서 m은 분포도 p와 q를 위한 양자화 단계이다. 두 개의 분포에 대한 거리는 식(3.10)과 같이 정의된다.

$$d(y) = \sqrt{1 - p[\hat{p}(y), \hat{q}]} \quad \cdots\cdots (3.10)$$

눈과 눈이 아닌 영역에 대한 명암 분포에 대한 특성은 짝수와 홀수 필드 두 영상에 의해 확실히 특성 지어진다. 그러므로, 어두운 동공과 밝은 동공 영상에 대응하는 각기 다른 특징을 갖는 2개의 확률분포가 나타난다. 눈의 특징 확률 분포가 표현된 m*m크기의 어두운 동공과 밝은 동공 영역에서 2차원 결합 히스토그램을 유도할 수 있다.

기존의 Kalman 필터 방식과 제안된 합성 추적 방식을 비교하여 표 3-1에 나타냈다. 실험 영상에서 뜬 눈과 감은 눈, 감추어진 눈에 대응하는 프레임 수를 계수하였다.

그림 3-4에 눈을 감거나 얼굴 방향에 의하여 밝은 동공이 사라진 경우를 나타냈다. 이 경우 차 영상에서 밝은 동공 없이 눈 영역의 색상 상태분포에 따라 눈을 추적하게 된다. 그림 3-4에 평균 이동 추적기에 의하여 추적된 눈 영역을 사각형으로 나타냈다.

표. 3-1 동공 추적 비교

Table. 3-1 Tracking statistics Comparsion for both pupil trackers

3000프레임 영상		Kalman 필터 추적기		제안된 합성 추적기	
		인식된 프레임 수	인식률	인식된 프레임 수	인식률
왼쪽 눈	(뜬눈) 2260 프레임	2012	89%	2255	99%
	(감은눈) 330 프레임	0	0%	317	95%
	(감추어진눈) 410프레임	0	0%	384	94%
오른쪽 눈	(뜬눈) 2125 프레임	1945	92%	2117	99%
	(감은눈) 330 프레임	0	0%	307	93%
	(감추어진눈) 545프레임	0	0%	499	92%
전체 인식률		3957	66%	5879	98%

그림. 3-4 평균이동 추적기를 이용한 동공 추적
Fig. 3-4 Tracking of using mean shift pupil tracker

3.4 시선 식별 전처리

본 논문에서는 2장에서 기술한 시선 식별 법 중 적외선을 이용하여 동공과 글린트를 추적한 방법을 개선하여 얼굴의 움직임이 있는 경우나 개개인에 대하여 교정 작업 없이 동작하는 시선 식별 시스템을 제안한다.

3.4.1 동공 검출

본 논문에서는 동공영상을 획득하기 위하여 Hutchinson의 방법에 근거하여 880nm파장의 적외선 LED를 두개의 고리 모양으로 구성한 능동적 적외선 조명기를 제작하여 40nm파장에서 32mW의 강도로 적외선을 조사한다.

이 적외선 광원은 사용자와의 거리에 관계없이 그림 3-5와 같이 조사 각도에 따라 밝은 동공 영상과 어두운 동공영상을 생성할 수가 있다. 다만 안경을 쓴 사용자의 경우 빛의 반사에 따른 오인식이 있을 수 있으므로 안경을 착용하지 않은 경우로 가정하고 카메라와 사용자의 거리는 사용자가 컴퓨터 앞에 자연스럽게 앉아 있을 경우인 0.5~1m로 한다.

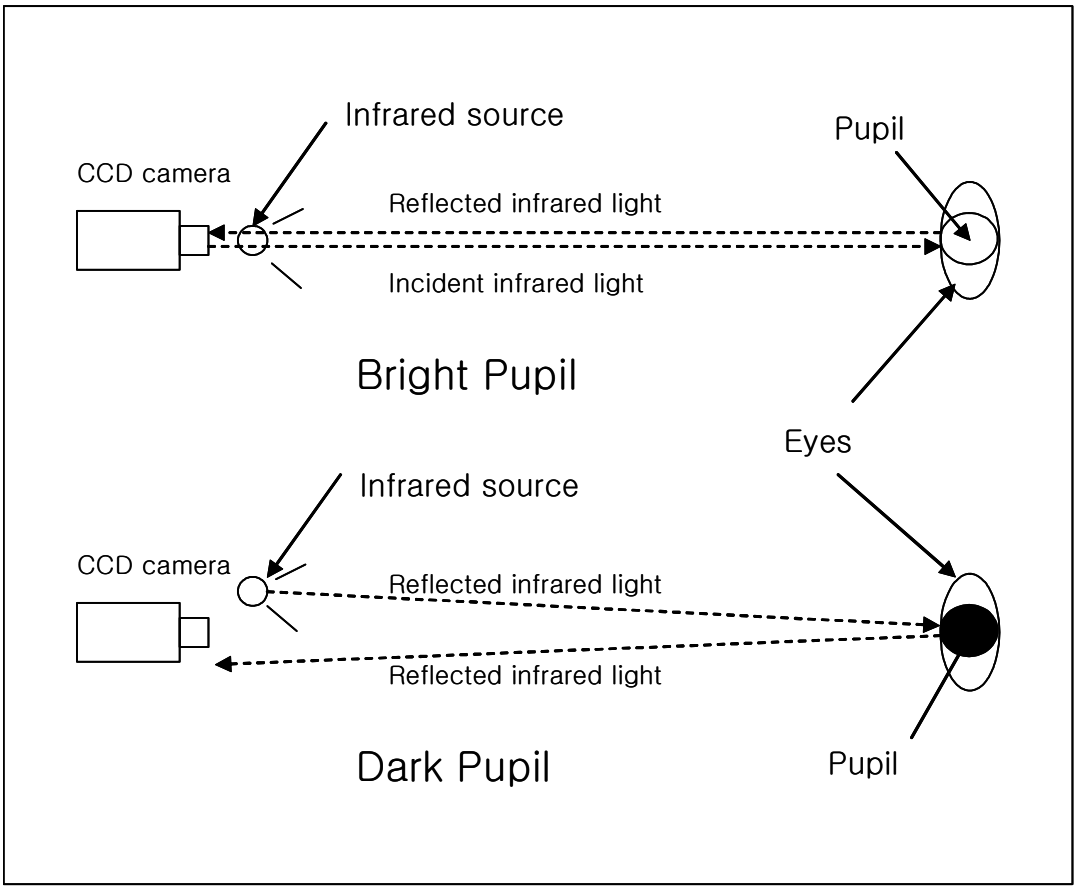


그림. 3-5 밝고 어두운 동공 효과 원리
Fig. 3-5 Principle of bright and dark pupil effect

광원을 교차 조사함에 따라 안쪽 고리의 적외선 LED들이 켜졌을 때는 동공에서 반사된 적외선이 검출되어 밝은 동공영상이 생성되고, 바깥 쪽 고리의 적외선 LED들이 켜졌을 때는 동공에 반사된 적외선이 검출되지 않으므로 어두운 동공이 생성된다. 적외선 조명 이외의 광원들로부터 간섭을 최소화하기 위하여, 10nm파장만을 통과시키는 광학 대역 필터를 사용함으로써 신호 대 잡음비를 20이상 증가시키어 획득한 영상의 질을 향상 시킨다. 그림 3-6은 중심이 같은 2개의 적외선 LED고리와 대역 필터로 이루어져 있는 적외선 조명기를 나타내고 있다.

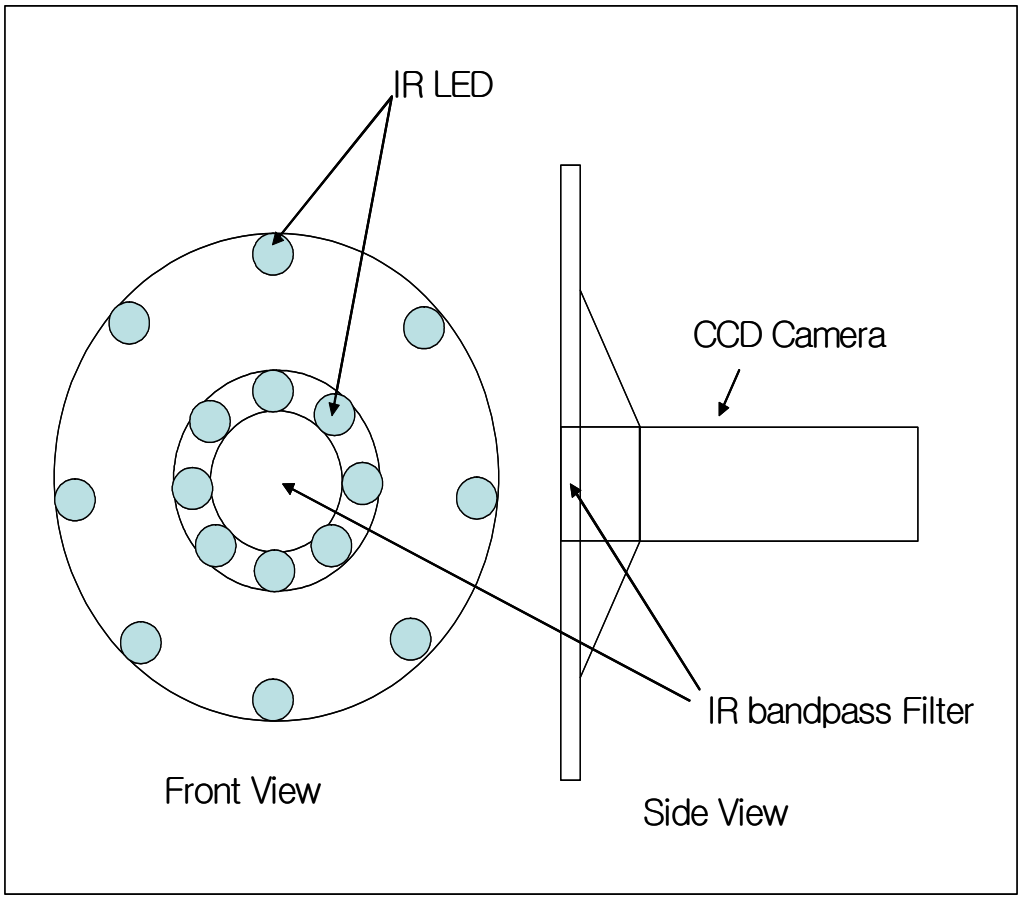


그림. 3-6 하드웨어 구성 : 능동적 적외선 조명 카메라

Fig. 3-6 Hardware Setup : the Camera with an Active IR illuminator

다양한 조명 조건과 여러 얼굴 방향에 대해 동공 검출 및 추적하는 것이 쉽지 않기 때문에 얼굴의 위치 설정 후 두 눈의 정확한 위치를 발견한다.

참조 점에 일치된 두 눈으로 인하여 눈으로부터 코와 입에 대한 상대적인 거리를 개인별로 허용오차 범위 내에서 고유하게 설정된다. 눈의 수평위치(eye line)를 설정하기 위해 주어진 영상의 수직 에지를 수평 라인별로 검색하였고, 에지가 있는 위치에 해당하는 그레이레벨을 누적함으로써 eyeline을 발견 할 수 있다. 발견된 eyeline을 따라 수직 line을 발견하기 위한 windowing을 하고 윈도우내의 수직 edge의 위치에 해당하는 그레이레벨을 누적한다. 누적 합이 가장 큰 윈도우의 중앙 열의 수직의 eyeline으로 설정할 수 있다.

그림 3-7 (a), (b)와 같은 짝수 필드 영상과 홀수 필드 영상을 만들어 같은 배경과 외부 조명을 사용하지만 짝수 필드의 동공이 홀수 필드보다 밝게 보였고, 배경과 외부조명의 간섭을 제거하고자 짝수 필드 영상에서 홀수 필드 영상을 제거하여 차 영상을 만들었다. 차 영상에서는 대부분의 배경과 외부 조명의 간섭효과가 제거되었다. 그림 3-7의 (c)에 차 영상을 보여주고 있다. 차 영상은 동공 영역 외에도 많은 다른 영역을 나타낸다.

차 영상은 동공 영역 외에도 많은 다른 영역을 나타낸다. 눈 영역을 찾아내기 위하여 우선적으로 동공이 있을 만한 모든 영역을 그림 3-8과 같이 표시하였다.

(a) 짝수 필드영상

(b) 홀수 필드 영상

(c) 차 영상

그림. 3-7 (a) 짝수필드영상 (b) 홀수필드영상 (c) 차 영상

Fig. 3-7(a) Even field image (b) Odd field image
(c) Differential image

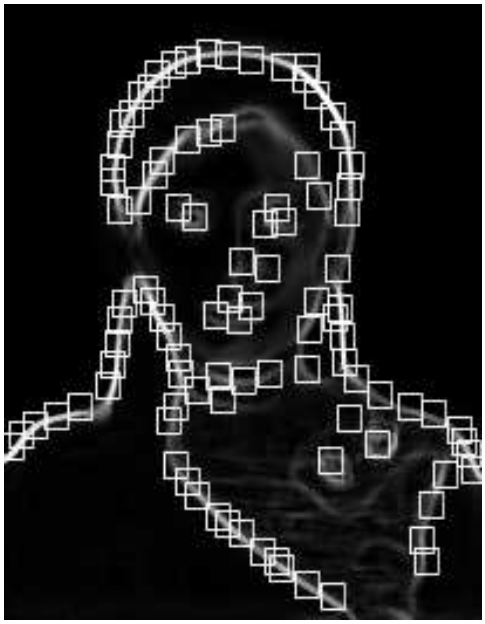


그림. 3-8 동공 후보 영역을 표시한 영상.
Fig. 3-8 The image marked by pupil candidates.

3.5 매개변수 추출

시선 식별은 동공과 글린트를 검출하고 추적하여 적합한 매개변수를 추출함으로써 시작된다. 시선 식별을 위해 밝은 동공 효과와 어두운 동공 효과를 만들어 내어 동공을 검출하고 추적할 수 있었다. 또한 매 프레임에서 동공의 위치는 중심 궤적에 의해서 특성 지어졌다. 각막에 반사되는 글린트는 양쪽 영상 모두에서 나타났으나, 밝은 동공 영상에서는 동공과 같은 위치에 겹쳐지게 되는 경우가 많이 발생하여, 어두운 동공 영상에서 보다 쉽게 검출 할 수 있었다. 즉, 어두운 영상에서 글린트는 다른 부분보다 밝게 나타나므로 글린트 검출과 추적을 수월하게 할 수 있었다. 동공 검출과 추적 방법은 어두운 영상에서 글린트를 검출하고 추적하는데도 사용하였다[45].

시선 식별을 위한 1차적인 매개변수로 글린트 좌표, 동공과 글린트와의 변화량과 기울기를 그림 3-9와 같이 구할 수 있다.

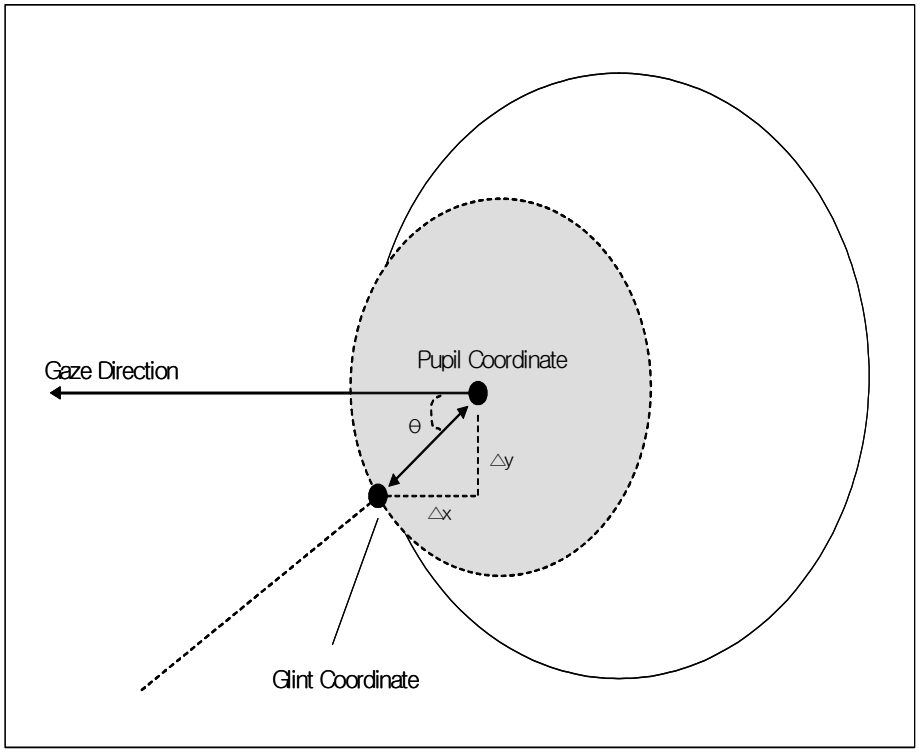


그림. 3-9 시선 식별을 위한 매개변수

Fig. 3-9 Parameter of gaze discrimination

시선 식별을 위하여 1차적인 매개변수만을 이용할 경우 움직임이 있는 경우에 대한 시선 식별이 잘 이루어지지 않는 경향이 있다. 본 논문에서는 얼굴 움직임에 대비한 얼굴 방향에 대한 매개변수를 추출한다. 이때, 얼굴방향에 대한 동공 특성은 다음과 같다.

얼굴이 정면으로부터 측면으로 회전함에 따라 동공 간 거리는 좁아진다.
두 동공의 평균 명암 비율은 얼굴이 좌우 또는 상하로 회전하는 동안 1보다 커지거나 1보다 작아진다.
두 동공의 모양은 얼굴이 좌우 또는 상하로 회전하면서 보다 타원형이 된다.
동공의 크기는 얼굴이 좌우 또는 상하로 회전하면서 작아진다.

위와 같은 사실을 바탕으로 동공으로부터 얼굴 방향을 추정하기 위한 기초로 사용하여, 좌우동공크기, 좌우동공간 거리 및 좌우 동공 모양과 같은 좌우 동공 명암을 이용하여 얼굴 방향을 분석 하여, 이와 같은 특징들이 일관성을 가질 수 있도록 이 매개변수들을 얼굴 전면에서의 측정값으로 나누어줌으로써 표준화하였다. 동공 특징공간은 7가지 특성에 의해 구축되어지므로 클러스터들은 더욱 뚜렷하게 구별되어 질 수 있으며, 얼굴방향은 얼굴 특징공간에서의 동공의 특성에 대한 투영을 통해 결정할 수 있다.

본 논문의 제안된 방법의 문제점은 동공이 눈꺼풀로 가려질 때 동공의 식별이 이루어지지 않을 수 있다는 것이다. 이러한 문제를 해결하기 위해서, 특정 시점에 대한 얼굴 방향을 그 전후 프레임들에 대한 얼굴 방향 측정치들을 포함한 평균치로 계산한다. 또한 머리의 움직임이 자연스럽다는 가정 하에, 전후 시점에 대해 측정된 얼굴 방향 등과 같은 경험적인 제약 조건이 필요하다. 그림 3-10은 얼굴방향 식별에 대한 알고리즘을 순서도로 나타내었다.

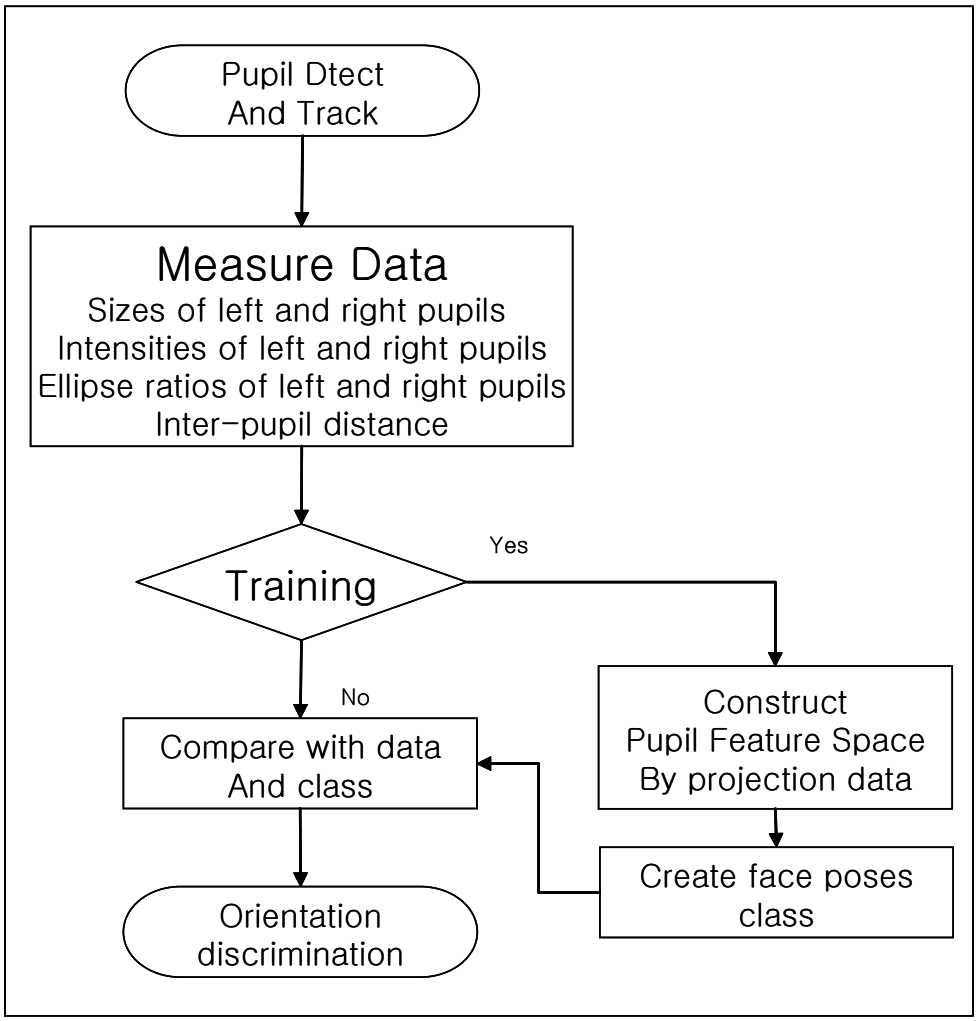


그림. 3-10 얼굴 방향 식별 알고리즘 순서도

Fig. 3-10 Face pose discrimination algorithm flowchart

3.5 시선교정

본 논문에서는 시선 식별을 위하여 전처리 과정에서 얻는 동공과 글린트를 가지고 동공-글린트 벡터를 구성하고, 동공-글린트 벡터를 모니터 상에 매핑하는 기능을 수행토록 하였다. 동공과 글린트에 관한 매개변수가 충분히 주어진다면 시선에 따라 유일한 기능을 갖는 동공과 글린트 매개변수를 찾고 GRNN을 이용하여 교정을 일반화 한다[46]. GRNN위상은 입력 층, 숨은 층, 합산 층. 그리고 출력 층으로 구성하였다. 입력 층은 1개의 출력 층에 대하여 6개의 입력을 갖도록 하였고, 숨은 층 노드들의 수는 학습 샘플의 수와 같게 하였다. 합산 층의 뉴런 수는 출력 노드 수에 1을 더한 것과 같게 하였다. 그림 3-11은 GRNN구조를 보여주고 있다.

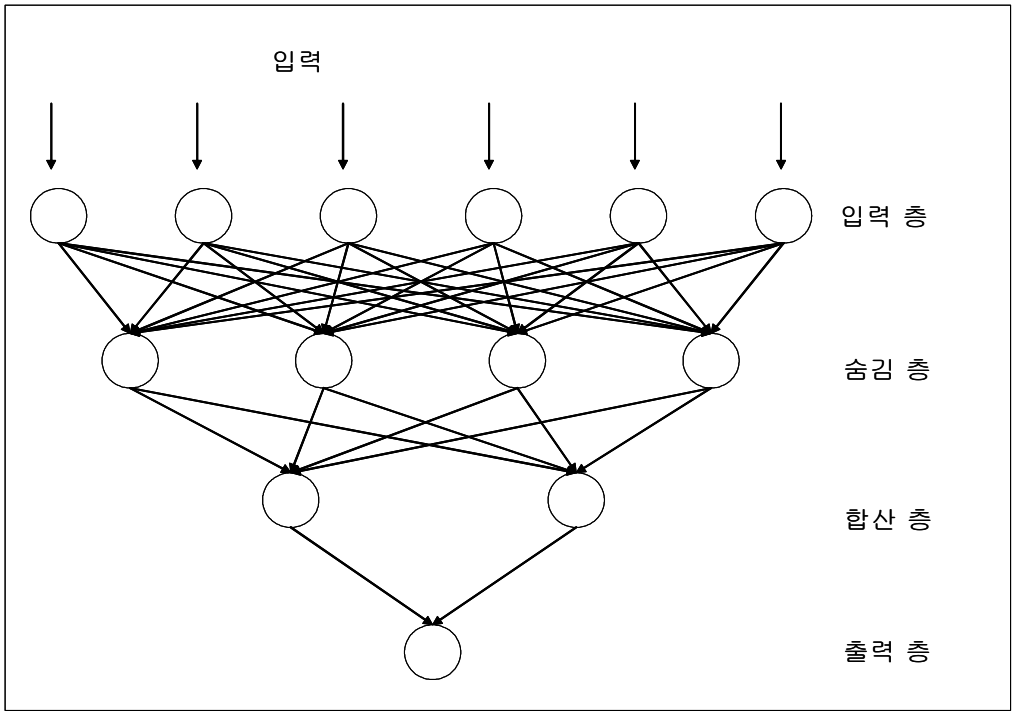


그림. 3-11 시선 교정을 위한 GRNN구조

Fig. 3-11 GRNN architecture used for gaze calibration.

수직 수평 공간적 시선 분석에 의하여 출력 노드에서 시선 좌표 각각 Sx와 Sy를 나타내는 두 개의 동일한 GRNN을 구성한다. 입력 층에 사용할 매개변수는 얼굴 거리와 방향에 따라 다양한 변화를 주어야 한다. 입력 벡터는 GRNN에서 사용되기 전에 적합하게 정규화 되어야 된다. 정규화는 모두 입력 특징들이 같은 범위에 있게 하여주며 다른 머리 위치에 따른 막대한 학습 데이터가 GRNN의 학습으로 모아진다. 학습데이터가 수집되는 동안에, 사용자가 각 시선 영역에 시선을 고정하도록 요청하여, 각각의 시선이 고정되는 동안 적합한 10개의 매개변수 조합을 수집하도록 하였다. GRNN에 사용할 입력 벡터는 다음과 같다[47].

$$g = \Delta x \, \Delta y \; r\theta \; g_x \, g_y \qquad (3.11)$$

여기서 Δx와 Δy는 동공과 글린트 사이의 변화이고, r은 동공에 맞는 타원의 비율이다. 그리고 θ는 타원의 기울기 이고, g_x와 g_y는 글린트 영상의 좌표이다.

입력 매개변수는 다음에 근거하여 선택 하였다. Δx와 Δy는 글리트와 동공사이의 상대 이동을 위하여 사용한다. 글린트-동공 벡터의 크기는 적외선 조명기에 내장된 카메라와 피사체의 거리와도 연관된다. r은 얼굴 방향을 계산하기 위하여 사용한다. 얼굴이 정면을 향하고 있다면 비율은 1에 가까울 것이고, 얼굴이 좌우나 위아래로 움직인다면 비율은 1보다 커지거나 작아진다. 기울기 θ는 얼굴 회전을 위하여 사용한다. $(g_x \, g_y)$는 평면 얼굴변환을 위하여 사용되어진다. 이러한 매개변수들을 사용함으로써 얼굴과 동공이 움직이면 얼굴 움직임과 동공 이동을 계산하여 그에 부합하는 매개 변수로 갱신함으로써, 머리 움직임에 따른 영향을 감소할 수 있다. 또한 매개변수들은 동공의 크기에 대해서 독립적이므로 시선 매핑 기능을 일반화 시킬 수 있어 다른 사용자를 위한 재계산 작업을 효과적으로 제거할 수 있었다.

IV. 실험 결과 검증 및 평가

4.1 얼굴 동작 분류에 대한 시선 식별

본 논문에서 제안된 방법은 특정한 분야에 응용이 가능하다. 운전자의 졸음 또는 피로 상태를 측정할 수 있는 잠재적인 시각정보는 사람의 얼굴 표정을 분석하여 이용하는 방법이 있다. 일반적으로 사람은 주의력 정도에 따라 서로 다른 얼굴 표정을 나타내는 경향이 있는데, 졸음 상태에 있는 사람의 얼굴 표정은 축 늘어진 얼굴 근육, 무표정 또는 하품 등에 의해서 특성화 될 수 있다. 특히 눈과 입주위의 얼굴 특징들은 얼굴 표정을 만드는데 아주 중요한 공간적 패턴들을 제공한다. 이러한 패턴들은 얼굴 표정을 특성화 하여 분석하는데 사용된다.

검출된 얼굴 특징 점은 각 요소의 움직임을 파악하게 되고 얼굴 표정을 분석하는데 사용되었다. 특히 가장 잘 구별되는 하품의 경우 입의 움직임을 관측함으로써 획득 할 수 있고 입을 다물고 있는 상태에 비하여 확연히 수직 방향으로 입술의 형태가 변한다. 윗입술과 아랫입술사이의 거리와 좌우 입술 모서리 영역으로부터 얻은 길이를 이용하여 입을 벌리고 있는지 닫고 있는지 알 수 있다.

총 5명의 피검자를 대상으로 각각 실내에 앉은 상태로 각성 혹은 부주의와 같은 상태와 하품, 잠듬과 같은 얼굴 눈꺼풀의 변화량에 따른 얼굴 시선 영역을 추정한다. 실험 환경은 적외선 카메라 시스템을 이용하여 초당 20프레임 320*240 해상도로 얼굴 영상을 약 1분씩 총 5회 기록하여 실험한다. 이때 같은 동작으로 모든 단계 부주의, 하품, 잠듬에 대하여 상향식 분류에 사전 확률을 설정한다.

그림 4-1은 얼굴 동작에 따른 분류 결과를 측정한 결과를 나타낸다. 실험을 위해 눈 검출 영역, 머리 방향추정, 머리 움직임추정, 얼굴추적 등을 조합하여 a, b, c, d, e, f 로 동작을 가정하고 각 동작으로 모든 단계에 부주의, 하품, 잠듬에 대해 사전 확률을 측정한다. Ia, Yw, Fa는 각각 부주의, 하품, 잠등을 나타낸다.

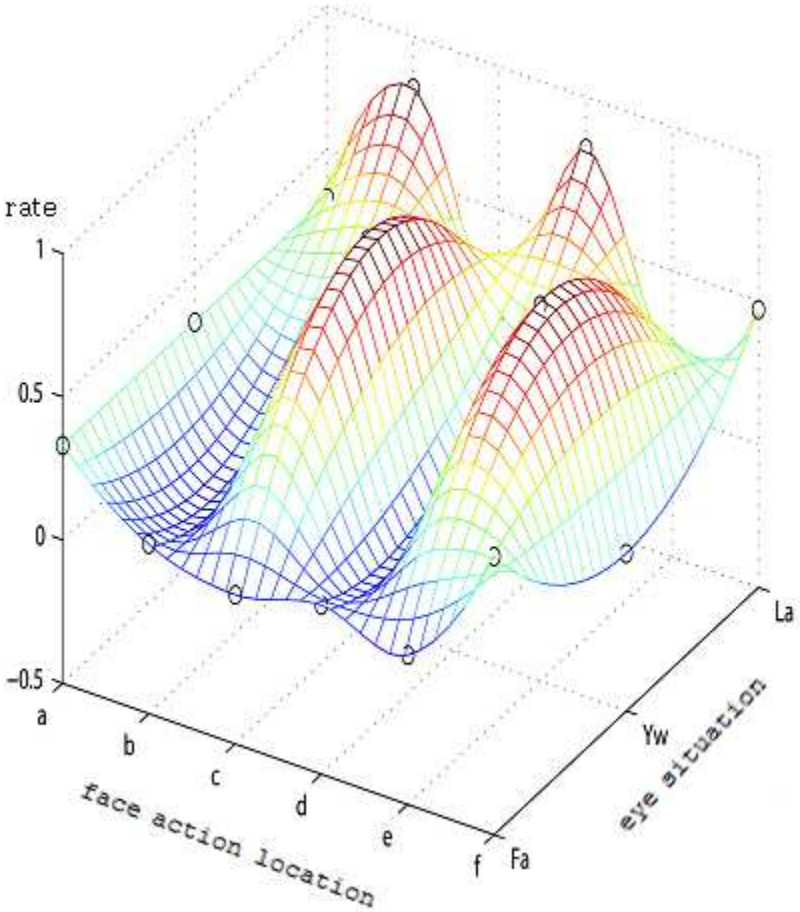


그림. 4-1 얼굴 동작 분류 결과

Fig. 4-1 The result of a face action grouping

그림 4-1에서 a는 Ia, Yw, Fa 값들이 0.5보다 작은 값인 0.33으로 동일하게 측정되었다. 이것은 각성 상태이거나 각성 상태일 경우 얼굴 동작의 변화가 없는 상태를 의미한다. b는 부주의 상태에 대한 높은 사전 확률을 나타내었다. c는 Yw가 0.83으로 높게 측정이 되어 하품에 해당되고 d와 e는 분류 결과가 각각 부주의와 하품으로 측정된다. 즉, a, b, c, d, e는 피검자들의 여러 동작에 따른 얼굴 영역에서 시선 검출이 용이하다. 그러나 f는 부주의와 잠듬으로 사전 확률이 0.47로 동일하게 측정되었다. 이 경우 얼굴 정보가 불분명 시에는 얼굴 후보영역에서 얼굴 특징 점에 따른 시선 검출에 어려움이 있다.

홍채 및 동공 영역을 추적하는 방법들은 일반 조명하에서 영상처리 방법에 의해 눈동자의 움직임을 추적하므로 특수 조명이 필요 없고 손쉽게 구현할 수 있다는 장점이 있는 반면, 동공과 홍채 사이의 계조도 차이가 아주 작기 때문에 정확한 추출이 어렵고 정확도가 낮은 단점이 있다. 이를 극복하기 위해 모니터 상의 시선 위치를 파악할 수 있도록 적외선 조명기를 구성하여 밝은 동공 효과를 얻어 동공을 실시간으로 검출, 추적하였다.

4.2 학습에 참여하지 않은 사용자의 얼굴 움직임이 있는 실험 결과

본 논문에서는 동공을 검출한 후 연속된 프레임에서 지속적으로 동공의 위치를 추적하기 위하여 베이지안 네트워크와 Kalman필터 알고리즘을 이용하여 동공 위치를 추적한 결과 98%의 추적 률을 나타내었다. 이후 제안된 식별 알고리즘을 실험하기 위해 모니터상의 스크린을 몇 개의 영역으로 분할하고, 사용자가 얼굴을 자연스럽게 움직이게 하면서 레이저 포인터로 지시하는 화면을 응시하도록 하였다. 그리고 각각의 프레임에서 시선을 실시간 식별하여 실제 시선 위치와의 정확도를 측정하였다. 모니터상의 스크린 영역은 머리 움직임과 같은 외부 요인에 따라 수직 동공-글린트 벡터 추정이 수평에 비해 보다 민감하게 작용함으로써 동공의 수직 이동범위가 수평 이동 범위보다 작게 나타나고, 수평 데이터에 비해 수직 데이터의 신호 대 잡음비가 낮게 됨으로써 시선의 수직 분류를 수평 분류보다 작게 잡아 모니터 영역을 2*4로 나누어, 얼굴 움직임이 있는 경우와 사용자에 따른 교정 작업 없이도 견실하게 동작하는지를 확인하기 위하여 학습에 참여하지 않은 사용자를 피검자로 하였다.

학습에 참여하지 않은 사용자에 대한 실험 결과는 표 4-1과 같이 전체적인 결과로 볼 때 약 84%정도의 식별 률을 나타내었다. 표 4-1에서 보는 것과 같이 맨 우측 영역인 4,5영역은 인식률이 98, 100%인 반면, 4,5영역과 인접한 3,6 영역은 인접한 영역으로 오인식된 경우가 많이 나타나 현격하게 저조하여 약 77%정도의 인식률을 나타내었다. 즉, 학습에 참여하지 않고 얼굴의 움직임이 있는 관계로, 일부 영역에서 인접한 영역으로 인식되는 경우가 많아져 학습에 참여한 경우보다 인식률이 현격하게 저하되는 것을 알 수 있으며, 실제 응시 영역을 바르게 식별을 못하고 인접한 영역으로 식별한 것을 확인 할 수 있다. 표 4-1에서 영역 1을 바라보는 시선의 18%가 오인식되어 영역 2로 나타났고, 영역 3을 바라보는 시선의 23%가 영역 4로 인식되었다. 결과적으로 오 인식은 인접한 시선영역에서 주로 나타나는 것을 알 수 있다.

표. 4-1 학습에 참여하지 않은 사용자의 얼굴 움직임이 있는 상태에서의 시선식별

Table. 4-1 Gaze discrimination results using a person whose data is not included in the training set and with head movement.

실제 주시 영역	데이터 수	시선 식별 결과								성공률
		1	2	3	4	5	6	7	8	
1	300	245	55	0	0	0	0	0	0	82%
2	300	0	260	40	0	0	0	0	0	87%
3	300	0	0	230	70	0	0	0	0	77%
4	300	0	0	0	295	5	0	0	0	98%
5	300	0	0	0	0	300	0	0	0	100%
6	300	0	0	0	30	40	230	0	0	77%
7	300	0	0	10	0	0	25	265	0	88%
8	300	20		0	0	0	0	30	250	84%
전체	2400	2075개의 시선을 바르게 식별								84%

4.3 인접 영역에 대한 재식별 인식률 결과

시선 식별에서 오 인식을 줄이기 위한 일련의 방편으로, 인식된 결과에 대하여 비교 대상을 줄여 다시 한번 식별을 하도록 하는 검증을 하였다. 시선 영역에서 오 인식을 확인하여 이 영역에 대한 재식별을 하도록 하는데 주안점을 두었다. 검증을 위한 재식별은 각각의 시선 영역에 인접한 영역과 다시 식별을 수행하도록 하는 것이다. 인접 영역에 대한 정의를 그림 4-2에서 보여주고 있다.

각각의 시선영역에 대한 인접 영역을 확인하고 인접한 영역에 대한 학습 데이터만을 이용하여 학습을 수행한 후, 인접한 영역만으로 재식별 하도록 하여 식별 결과를 검증한다. 즉, 식별 결과 영역 1로 식별되었을 경우, 검증을 위하여 영역 1,2,8의 학습데이터만을 가지고 재학습하여, 영역 1로 1차 식별한 결과가 인접한 영역인 2,8영역을 혹 1영역으로 오인식한 것이 아닌지를 검증하도록 한 것이다.

기존에는 입력 벡터가 주어지면 식별기는 8개의 시선영역 중 하나의 시선 영역으로 식별하였고, 재식별기에서는 식별결과에 대응하는 인접영역의 학습 데이터만을 이용하여 재식별을 하도록 한 것이다. 즉, 만약에 입력벡터에 대하여 8개의 영역 1로 식별하였다면 그 결과에 대하여 영역 1,2,8의 학습 데이터만을 가지고 재식별 토록 하여 최종 출력을 얻도록 한 것이다.

재식별을 수행한 결과를 표 4-2에 나타내었고, 그림 4-3에 각 영역별 시선식별 결과를 도식화 하였다. 시선 영역에 대한 인접 영역의 재식별을 한 경우가 하지 않은 경우보다 9%정도 정확성이 향상되어 약 93% 정확도를 얻을 수 있었다. 특히, 인접영역 1과 2사이의 오 인식률은 18%에서 8%로, 시선 영역 3과 4사이에 오 인식률은 이전 24%에서 5%로 감소하였다. 다른 시선 영역에 의한 오 인식 에러 역시 향상되었거나 동일한 성능을 나타내었다. 그리고 화면을 4*2로 양자화 하여 실험한 결과 공간 시선 해상도는 화면에서 1m정도 떨어진 거리에서 수평으로 10cm, 수직으로 약 13cm정도였다. 즉, 수평으로 8도 정도 수직으로 5도정도 나타났다.

제안된 시선 식별기의 해상도가 기존의 다른 시선 식별기보다 높다고 할 수는 없지만, GRNN을 통하여 시선 교정 작업을 일반화함으로써 개개인에 대한 교정 작업이 필요 없고, 얼굴의 움직임이 있는 경우에도 견실하게 동작하는 것이 큰 이점이라 할 것이다.

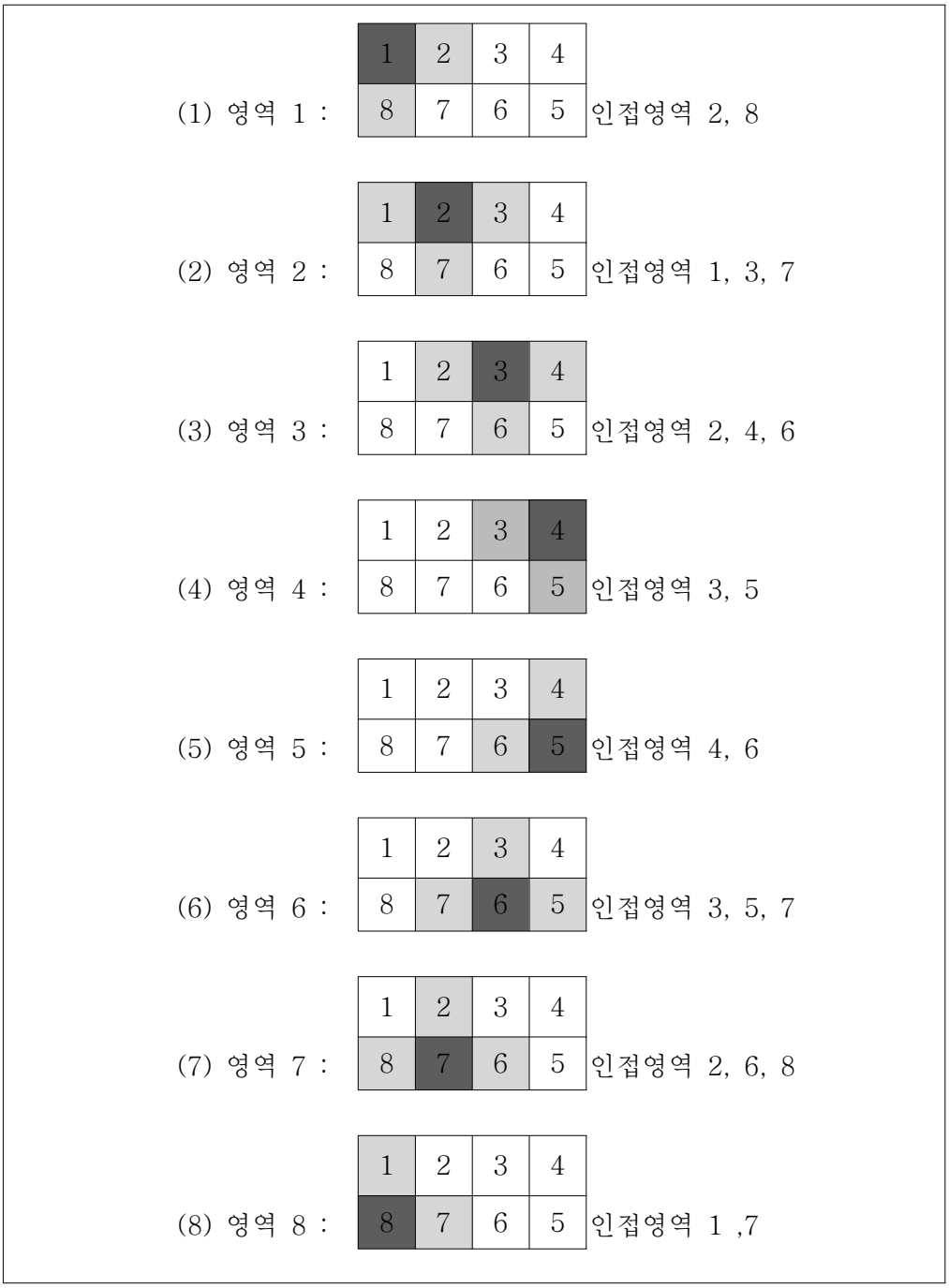


그림. 4-2 시선영역에 대한 인접 영역

Fig. 4-2 Neighboring gaze classes for each gaze region

표. 4-2 재식별기를 이용한 시선 식별 결과

Table. 4-2 The gaze discrimination results using the gaze reclassifier

실제 주시 영역	데이터 수	시선 식별 결과								성공률
		1	2	3	4	5	6	7	8	
1	300	275	25	0	0	0	0	0	0	92%
2	300	0	290	10	0	0	0	0	0	97%
3	300	0	0	285	15	0	0	0	0	95%
4	300	0	0	0	295	5	0	0	0	98%
5	300	0	0	0	0	300	0	0	0	100%
6	300	0	0	5	25	25	245	0	0	82%
7	300	0	0	10	0	0	25	265	0	88%
8	300	15		0	0	0	0	10	275	92%
전체	2400	2230개의 시선을 바르게 식별								93%

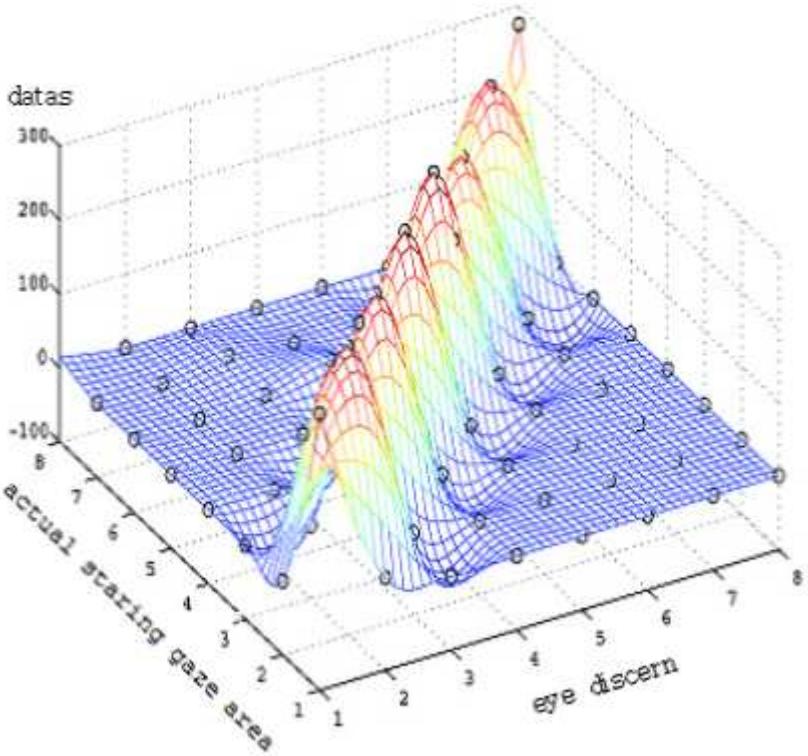


그림. 4-3 재식별기를 이용한 시선 식별 결과

Fig. 4-3 The gaze discrimination results using the gaze reclassifier

V. 결　　론

기존의 시선식별 연구는 얼굴의 움직임을 이용하는 방법과, 눈동자의 움직임을 이용하는 두 가지 방법이 있다. 첫 번째 방법은 얼굴의 위치에 따라 시선의 위치를 결정하되 미세한 시선 변화를 감지하지 못한다는 단점을 가지고 있다. 눈과 시선식별 시스템의 또 다른 문제는 각각의 사용자에 대하여 교정 작업이 필요하다.

따라서 이와 같은 기존의 문제점 해결을 위해 본 논문에서는 Kalman 필터 알고리즘을 적용하여 현재 머리의 위치 정보를 이용하여 미래 위치를 추정하였고 머리 영역 검출 후 적절한 상태로 판명되면 추출된 후보 영역이 얼굴인지의 여부를 판단하였다. 또한 얼굴의 진위 여부를 판단하기 위해서 얼굴의 특징요소인 양쪽 눈, 양쪽 콧구멍, 입의 양끝을 사용하여 구조적 정보와 비교적 처리시간이 빠른 수평, 수직 히스토그램 분석법을 이용하여 얼굴의 요소를 검출하였다.

여기서, 눈의 위치는 전체 얼굴 크기의 1/2정도 높이, 코끝의 위치는 눈 위치의 1/2정도 높이에 위치하고, 입은 코의 위치로부터 1/2정도 되는 높이에 위치하도록 하였다. 이러한 구조적 정보를 기반으로 얼굴 요소의 후보 위치를 선정하고 그 위치 내에서 얼굴 요소를 추출하였다. 얼굴 영역에 대한 이진 영상(binary image)을 구한 다음 얼굴의 구조적 정보를 기반으로 해 눈이 존재한 것으로 예측되는 영역에서 수직, 수평 투영 히스토그램을 구하였다. 눈이 존재할 경우는 수평축과 수직축에서 히스토그램의 피크 위치가 나타나고 이 축이 만나는 두 점이 양 눈의 위치가 되도록 하였다.

홍채 및 동공 영역을 추적하는 방법들은 일반 조명하에서 영상처리 방법에 의해 눈동자의 움직임을 추적하므로 특수 조명이 필요 없고 손쉽게 구현할 수 있다는 장점이 있는 반면, 동공과 홍채 사이의 계조도 차이가 아주 작기 때문에 정확한 추출이 어렵고 정확도가 낮은 단점이 있다. 이를 극복하기 위해 모니터 상의 시선 위치를 파악할 수 있도록 적외선 조명기를 구성하여 밝은 동공 효과

를 얻어 동공을 실시간으로 검출, 추적하였고 동공-글린트 벡터를 추출할 수 있도록 하였다.

그리고 교정 작업 없이 얼굴 움직임이 있는 경우에 정확하게 동작하는 시선 식별 시스템을 제안하기 위해서 모니터 좌표를 4 * 2로 양자화 하여 실험한 결과, 학습에 참여하지 않은 사용자가 얼굴의 움직임이 있을 경우에 2400개의 프레임 중 2075개의 프레임에 대하여 바르게 시선을 식별하여 약 84%의 정확성을 얻을 수 있었다. 그러나 바르게 인식되지 않은 결과를 분석한 결과 인접한 영역에 대하여 어느 정도 오·인식이 있음을 알 수 있었다. 따라서 이와 같은 오·인식을 줄이고자 식별 결과에 대해 재식별을 하여 시선 영역을 검증토록 하였다.

연구에서 인접영역의 시선을 재식별 하였을 경우 1과 2사이의 오 인식률은 18%에서 8%로, 3과 4사이에 오 인식률은 24%에서 5%로 감소하였다. 기존의 84% 정확성에 비해 약 9%정도 정확성이 향상되어 93%를 나타내었다. 제안된 시선 식별기는 자연스러운 얼굴 움직임이 있는 경우나 사용자에 따른 교정 작업이 없음에도 실시간 시선식별의 성능이 향상되었다.

향후, 사용자의 안경 착용과 같은 제한 사항을 극복하고 시선의 공간 해상도를 보다 향상시키고 아울러 시선 영역을 좀 더 세분화 한 연구가 필요하다.

참고문헌

[1] J.Weng, D. L. Swets, "Face Recognition in Biometrics : Personal Identification in Networked Society", pp.67-86, Boston. MA KluwerAcademic, 1999.

[2] K. Sung, T.Poggio, "Example-based learning for view-based human face detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.20 Issue:1, pp.39-51, Jan. 1999.

[3] L. Cohen, N. Sebe, A. Garg, M.A. Lew, T.S. Huang, "Facial expression recognition form video sequrences", IEEE International Conference on Multimedia

[4] T. Yamaguchi, M.Tominage, K, Mrakami, H. Koshimizu, "Regeneration of facial image eye-contacting with partner on TV Confernce environment", IEEE Iternational Conference on Systems, Man, and Cybernetics, vol.2, pp.1169-1174, 2000.

[5] C. Collet, A. Finkel, R. Gherbi, "CapRe; a gaze tracking system in man-machine interaction", IEEE International Conference on intelligent Engineering Systems, pp.577-581, 1997.

[6] P.Ekman and W.V. Friesen. "Emotion in the human face System." cambridge University Press, San Francisco, CA, second edition, 1998.

[7] State of the art in facial Animation 22, acm siggraph '89 course notes,1999.

[8] Isfan A. Essa and Alex P. Pentland. Facial expression recognition using visually extracted facial action parameters. In Int'l Work-shop on Automatic Face- and Gesture-Recognition '95, pages 35-40,1999.

[9] 윤영지, 나상동 “베이지안 네트워크에서 기하학적관계를 이용한 얼굴 동작 인식”, 한국해양정보통신학회, pp477-480, 2007.

[10] 송윤석, “로봇의 효과적인 서비스를 위해 베이지안 네트워크 기반의 실내 환경의 가려진 물체 추론” 정보과학회논문지 제12권 제1호, 2006.

[11] J. Luo, et al., "A Bayesian network-based framework for semantic image understanding", Pattern Recognition, vol. 38, pp.919-934, 2005.

[12] L. D harmon, "Identification of human face profiles by computer", Pattern Recognition, pp.301-312. 1999.

[13] K. Wong. H Law, and P. Tsang,"A system for recognition human face", Proc, ICASSP, 99.1638-1642. 2003.

[14] L. Sirovich and M. Kirby "Low-dimensional procedure for the characterization of human face", J. Opt. Society Am, pp.519-524, 2000.

[15] M. Nixon "Eye spacing measurement for facial recognition", SPIE Proc, pp.575:279-285, 2002.

[16] A. Yuille, D.Cohen, and P.Hallinan "Feature extrantion from face using deformable templates", IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.1:104-109, 2005.

[17] P. W. Hallinan, "Recognizing human eyes", SPIE Proc., pp214-226, 1991.

[18] B. S. Manjunath, R. Chellappa and C. Malshburg "A feature based approach to face recognition", IEEE Computer Society Conf. on Computer Vision and Pattern Recogniton, pp.1:373-378, 1992.

[19] J. M. Vicent "Facial feature location in resolution image by multilayered perceptrons", Artificial Neural Networks, : North-Holland, pp.821-826, 2002.

[20] H. Li. P Roviainen, R. Forchheimer, "3-D motion estimation in model-based facial image coding", IEEE Trans. on Pattern Analysis and Machine Intelligence, pp.545-555,1993.

[21] G. Yang and T.S. Huang, "Human face detection in a scene", IEEE

Conf, on Computer Vision and Pattern Recognition, pp.453-458, 1993.

[22] T. Fukuhara, T. Murakami, "3D-motion estimation of human head for model-based image coding", IEEE Proc., Vol.140, No.1, pp.26-35, 1993.

[23] Cheng. J, and Greiner. R, "Comparing Bayesian network classifiers.", Proveedings of UAI-99, 1999.

[24] Acid. S, and Campos, L. M., "An Algorithm for Finding Minimum d-Separating Sets in Belief Networks.", Proceedings of UAI-96, pp.3-10, 1996.

[25] Kevin Patrick Merphy, "A Brief Introduction to Graphical Models and Bayesian Networks.", Technical Report, Department of Computer Science, UC Berkley, 2000.

[26] Scott, D. & Findlay, J . M., "Visual search, eye movements and display units", Human factors report, University of Durkheim, United Kingdom, 1993.

[27] K. N. Kim, R. S. Ramakrishna "Vision-Based Eye-Gaze Tracking for Human Computer interface," Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp.117-120, Tokyo Japan, October, 1999.

[28] Baluja, S. & Pomerleau, D.A. "Non-Intrusive Gaze Tracking Using Artificial Neural Networks," Technical Report(CMU-CS-94-102) of School of Computer Science, Carnegie Mellon University, Pittsburgh PA, USA, 1994.

[29] Eye Gaze Monitoring (http://www.is.cs.cmu.edu/ISL.Multimodal.eye.html)

[30] X. Xie, R. Sudhakar, "On Improving Eye Feature Extraction using Deformable Templates", Pattern Recognition, Vol.27, No.6,

pp.791-799, 1994.

[31] J. Y. Deng, Feipei Lai, "Region-Based Template Deformation and Masking for Eye-Feature Extraction and description", Pattern Recognition, Vol.30, No.3, 99.403-419, 2001.

[32] Arrington Research ViewPoint EyeTracker (http://www.swling.net/~kfa/)

[33] Starker, I. & Bolt, R. A., "A gaze-responsive self-disclosing display", in Proc. of CHI '90, ACM, pp.3-9, 2003.

[34] 박호식, “GRNN을 이용한 실시간 시선 식별 시스템에 관한 연구”, 관동대학교 대학원 전자통신공학과 박사논문, 2005.

[35] P. Ballard et al., "Controlling a Computer via Facial Aspect". IEEE Trans. on SMC, vol. 25, no. 4, pp. 669-677, 1995.

[36] A. Gee et al., "Fast Visual Tracking by Temporal Consensus", Image and Vision Computing. vol. 14, pp.105-114, 1996.

[37] J. Heinz,ann et al., "3D Facial Pose and Gaze Point Estimation using a Robust Real-Time Tracking Paradigm". in Proc. of ICAFGR, pp. 142-147, 1998.

[38] T. Rikert et al., "Gaze Estimation using Morphable Models". in Proc. of ICAFGR, pp. 436-441, 1998.

[39] A. TOMONO et al., "Eye Tracking Method Using an Image Pickup Apparatus". European Patent Specification-94101635, 1994.

[40] Qiang Ji and Xiaojie Yang, Real time visual cues extraction for monitoring diver vigilance, International Workshop on Computer Vision Systems, July 7-8, 2001, Vancouver, Canada.

[41] T. E. Hutchinson, K. White, J. R. Worthy, n. Martin, C. Kelly, R. versus Templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, No.10, pp.1042-1052, Oct. 1993.

[42] T. E. Hutchinson, "Eye movement detection with improved calivration and speed", United States Patent,(4,950,069), 1998.

[43] K. OHMURA et al., "Pointing Operation Using Detection of Face Direction from a Single View". IEICE Trans. on Information and Systems, vol. J72-D-II, no. 9, pp.1441-1447, 2000.

[44] 나상동,박동석, 윤영지 “유전자 네트워크에서 확률적 그래프 모델을 이용한 정보 네트워크 추론”, 한국해양정보통신학회, 제10권, 제8호, 2006.

[45] 나상동, “GRNN 계층적 구조를 이용한 실시간 시선 식별 시스템 구현”, 한국통신학회, vol. 35, pp335-339, 2007.

[46] F. Hara and H.Kobayashi. Use of face robot for human-computer communication. In IEEE Conf. on SMC '95, pp 1515-1520, 1999.

[47] 박호식, 배철수, “능동적 적외선 조명을 이용한 실시간 3차원 얼굴 방향 식별”, 한국해양정보통신학회 제8권, 제3호, 2004.

ABSTRACT

A Study on real Time Gaze Discrimination Using Kalman Filter of Bayesian Network

Young-Ji Youn
Advisor : Prof. Sang Dong Ra, Ph.D.
Dept. of Information & Computer Science Education
Graduate School of Education, Chosun University

Eye-gaze is an input mode which has the potential of an efficient computer interface. Eye-gaze has been the focus of research in this area. Non-intrusive eye-gaze tracking that allows slight head movement is addressed in this paper. A small 2D mark is employed as a reference to compensate for this movement. The iris center has been chosen for purposes of measuring eye movement. The gaze point is estimated after acquiring the eye movement data. Preliminary experimental results are given through a screen pointing application.

Humans easily recognize where another person is looking at and often use this information for interspeaker coordination. We present a method based on three neural networks of th local linear map type which enables a computer to identify the head orientation of a user by learning from examples. One network is used for color segmentation, a second for localization of the face, and the third for the final recognition of the head orientation. The system works at a frame rate of one image per second on

a common workstation. We analyze the accuracy achieved at different processing steps and discuss the usability of the approach in the context of a visual human-machine interface.

The movement of user's eyes can provide a conventient, natural and high-bandwidth source of input. By tracking the direction of gaze of the user, the bandwidth of communcaiton from the user to the computer can be increased by using the information about what the user is looking at, and even designing objects specially intended for the user to look at.

A variety of eye-gaze tracking techniques have been reported in the literature. A short list includes Electro-Oculography, Limbus, Pupil and Eyelid Tracking Contact. Lens Method, Corneal and Pupil Reflection Relationship, Purkinje Image Tracking, Artificial Neural Networks and Head Movement Measurement.

Computer vision is intrinsically Non-intrusive, and does not require any overly expensive equipment. Non obtrusive sensing technology - such as video cameras and microphones - has received special attention in this regard. This paper draws on computer vision and image processing techniques for measuring eye-gaze.

저작물 이용 허락서

학 과	정보·컴퓨터 교육	학 번	20068116	과 정	석 사
성 명	한글: 윤영지 한문: 尹 泳 智 영문: Youn - Young - Ji				
주 소	광주광역시 북구 신안동 1-23				
연락처	HP : 010-3629-2601 E-MAIL : pardoner@nate.com				
논문제목	한글 : 베이지안 네트워크에서 Kalman필터 알고리즘을 이용한 실시간 시선 식별에 관한 연구 영문: A Study on real time Gaze Discimination Using Kalman Fillter Algorithm of Bayesian Network				

본인이 저작한 위의 저작물에 대하여 다음과 같은 조건 아래 조선대학교가 저작물을 이용할 수 있도록 허락하고 동의합니다.

- 다 음 -

1. 저작물의 DB구축 및 인터넷을 포함한 정보통신망에의 공개를 위한 저작물의 복제, 기억 장치에의 저장, 전송 등을 허락함
2. 위의 목적을 위하여 필요한 범위 내에서의 편집·형식상의 변경을 허락함. 다만, 저작물의 내용변경은 금지함.
3. 배포·전송된 저작물의 영리적 목적을 위한 복제, 저장, 전송 등은 금지함.
4. 저작물에 대한 이용기간은 5년으로 하고, 기간종료 3개월 이내에 별도의 의사표시가 없을 경우에는 저작물의 이용기간을 계속 연장함.
5. 해당 저작물의 저작권을 타인에게 양도하거나 또는 출판을 허락을 하였을 경우에는 1개월 이내에 대학에 이를 통보함.
6. 조선대학교는 저작물의 이용허락 이후 해당 저작물로 인하여 발생하는 타인에 의한 권리 침해에 대하여 일체의 법적 책임을 지지 않음
7. 소속대학의 협정기관에 저작물의 제공 및 인터넷 등 정보통신망을 이용한 저작물의 전송·출력을 허락함.

2009년 2 월 일

저작자 : 윤 영 지 (서명 또는 인)

조선대학교 총장 귀하