**Starburst : 특징 기반 및 모델 기반 접근법을 결합한 비디오 기반 눈 추적을위한 하이브리드 알고리즘**

Dongheng Li, David Winfield, Derrick J. Parkhurst 인간 컴퓨터 상호 작용 프로그램 Iowa State University, Ames, Iowa, 50010

**요약** *사용자의 시선을 알기. 안구 운동이 사용자의주의 집중 상태의 지표로 사용될 수 있다면 현재의 인간 - 컴퓨터 인터페이스를 향상시킬 수있는 중요한 잠재력이 있습니다. 안구 운동을 오늘날의 인터페이스로 통합하는 주요 장애물은 신뢰할 수 있고 저렴한 비용의 오픈 소스 안구 추적 시스템을 사용할 수 있다는 것입니다. 이러한 시스템을 인터페이스 설계자가 사용할 수있게 만들기 위해 우리는 피쳐 기반 및 모델 기반 접근 방식을 통합하고 오픈 소스 패키지에서 사용할 수 있도록 만든 하이브리드 아이 추적 알고리즘을 개발했습니다. 우리는 동공 피처가 감지되는 새로운 방식 때문에이 알고리즘을 "starburst"라고 부릅니다. 이 starburst 알고리즘은 순수 피쳐 기반 접근법보다 더 정확하지만 순수 모델 기반 접근법보다 훨씬 적은 시간이 소요됩니다. 현재 구현은 저렴한 헤드 장착 안구 추적 시스템에서 얻은 적외선 비디오의 안구 운동 추적에 맞춰져 있습니다. 유효성 확인 연구가 수행되었고이 기술은 약 1 도의 시야각의 정확도로 안구 위치를 안정적으로 추정 할 수 있음을 보여주었습니다.*

**1. 서 론** 안구 추적의 사용은 일상적인 인간 - 컴퓨터 인터페이스의 질을 향상시키는 중요한 잠재력을 가지고있다. 두 가지 유형의 휴먼 - 컴퓨터 인터페이스는 안구 운동 측정 (능동 및 수동 인터페이스)을 사용합니다. 액티브 인터페이스는 사용자가 안구 운동을 사용하여 명시 적으로 인터페이스를 제어 할 수 있도록 해줍니다 [8]. 예를 들어, 안구 타이핑은 전통적인 키보드처럼 수동으로 키를 누르는 대신 입력하기 위해 가상 키보드의 키를 봅니다 [9]. 이러한 활성 인터페이스는 이동 장애가있는 사용자가 컴퓨터와 상호 작용하도록 돕는 데있어 매우 효과적입니다. 이러한 기술은 사용자가 그래픽 사용자 인터페이스에서 아이콘을 선택하려고 할 때 일반적으로 아이콘을 먼저 보게되므로 눈을 추적 할 때 속도가 빨라질 수 있으므로 일반적인 인터페이스 사용에 유용 할 수 있습니다 [16]. 반면에 수동 인터페이스는 사용자의 안구 움직임을 모니터링하고 자동으로 사용자에게 적응합니다. 예를 들어, 비디오 전송 및 가상 현실 애플리케이션에서 시선을 끌 수있는 가변 해상도 디스플레이 기술

그림 1 : (a)및 (b) (c) 헤드 장착형 눈 추적 장치 (d) (c)장면 이미지 눈으로 얻은트래커. (d) 사용자의 오른쪽 눈이 적외선으로 조명 된 이미지. 명확하게 정의 된 어두운 동공과 적외선 LED의 반사를주의하십시오. 또한 소비자 등급의 기성품을 기반으로하는 저가형 건축물로 인해 (c) 및 (d)에 나오는 라인 소음의 정도를 확인하십시오.

niques는 시청자의 눈을 적극적으로 추적하고 시선을 사로 잡는 지점에서 높은 수준의 세부 사항을 제시하면서 산만하지 않은 주변부의 세부 수준을 희생합니다 [13, 14].

눈 추적은 많은 연구 시스템과 일부 소규모 소비자 제품에 배포되었지만 눈 추적은 그 잠재력에 도달하지 못했습니다. 눈 추적 기술은 다양한 방법 (예 : Purkinje 반사 기반, 콘택트 렌즈 기반 눈 코일 시스템, 전자 현미경 검사, 고전적인 눈 추적 기술 설문 조사 [19] 참조)을 사용하여 수년 동안 사용할 수있었습니다. . 이러한 기술을 휴먼 - 컴퓨터 인터페이스에 통합하기위한 주된 장애는 일상적인 사용에 너무 침해 적이거나 너무 비싸다는 것입니다. 최근, 헤드 장착 비디오 기반 아이 추적기의 소형화가 진행되면서 안구 추적의 침습성이 현저하게 감소하고있다 [15, 1]. 원격 비디오 기반의 안구 추적 기술은 또한 침입을 최소화하지만 [6, 10], 머리 장착 시스템에 대한 정확도가 떨어질 수 있습니다. 이러한 진보를 감안할 때 가장 중요한 장애물은 비용입니다. 현재에서 다수의 안구 추적기를 사용할 수

1

(a) (b)

시장있으며 가격은 약 5,000 ~ 40,000 US 달러입니다. 특히 고품질 디지털 카메라 기술의 가격이 지난 10 년 동안 급격히 떨어졌기 때문에이 비용의 대부분은 하드웨어 때문이 아닙니다. 오히려 비용은 맞춤형 소프트웨어 구현과 관련이 있으며 때로는 전문 디지털 프로세서와 통합되어 고속 성능을 얻습니다.

이 분석은 일상의 인간 - 컴퓨터 인터페이스에 안구 추적을 통합하기 위해 일반 컴퓨팅 하드웨어에서 실행되는 널리 사용 가능하고 신뢰할 수있는 고속 안구 추적 알고리즘의 개발이 구현되어야 함을 분명하게 나타냅니다. 이 목표를 향해 우리는 피쳐 기반 및 모델 기반 접근 방식을 통합하고 구현을 오픈 소스 패키지로 배포 할 수 있도록하는 하이브리드 안구 추적 알고리즘을 개발했습니다. 저비용 헤드 장착 안구 추적 시스템 [18]과 함께 안구 추적이 차세대 인간 - 컴퓨터 인터페이스에 성공적으로 통합 될 수 있다는 중요한 잠재력이 있습니다.

**2. 문제 성명** 위에서 언급했듯이 안구 추적 시스템은 원격 및 머리 장착 시스템으로 나눌 수 있습니다. 각 유형의 시스템에는 각각의 장점이 있습니다. 예를 들어 원격 시스템은 방해가되지 않지만 헤드 장착 시스템만큼 정확하거나 유연하지 않습니다. 다른 연구에서, 우리는 저가의 헤드 마운트 아이 추적기를 개발했다 [18]. 이 안구 추적기는 한 쌍의 안전 안경에 장착 된 2 개의 소비자 용 CCD 카메라로 구성됩니다 (그림 1 참조). 한 카메라는 눈의 이미지를 캡처하고 다른 카메라는 이미지의 이미지를 캡처합니다. 두 대의 카메라는 동기화되어 있으며 각각 640x480 픽셀을 캡처하는 30hz에서 작동합니다. 이 논문에서는 이러한 종류의 머리 장착 시스템에서 캡처 한 이미지에 사용할 수있는 눈 추적 알고리즘을 개발합니다. 그러나, 제안 된 알고리즘은 원격 시스템으로 캡쳐 된 비디오에도 적용될 수있다.

2 가지 유형의 이미징 프로세스가 눈 추적, 가시 광선 및 적외선 스펙트럼 이미징에서 일반적으로 사용됩니다 [5]. 가시 스펙트럼 이미징은 눈에서 반사되는 주변 광선을 포착하는 수동적 인 접근법입니다. 이러한 이미지에서 가시 스펙트럼 이미지에서 추적 할 수있는 가장 좋은 특징은 홍채와 limbus로 알려진 공막 사이의 윤곽입니다. 눈의 세 가지 가장 관련성있는 특징은 눈동자에 빛이 들어오는 구멍, 눈동자는 눈동자의 직경을 조절하는 색깔의 근육 그룹, 그리고 눈의 나머지 부분을 덮는 공막 인 공막입니다. 가시 스펙트럼 눈 추적은 제어되지 않은 주변 광이 소스로 사용되며, 여러 개의 반사 및 확산 구성 요소를 포함 할 수 있기 때문에 복잡합니다. 적외선 영상은 사용자가 감지 할 수없는 균일하고 제어 된 적외선으로 눈을 적극적으로 조명함으로써 제어되지 않는 반사를 제거합니다.의 또 다른 이점

적외선 영상은 윤곽보다는 동공이 영상에서 가장 강한 피쳐 윤곽이라는 것입니다 (그림 1d 참조). 공막과 홍채는 모두 적외선을 강하게 반사하고 공막 만이 가시 광선을 강하게 반사합니다. 동공 윤곽을 추적하는 것은 동공 윤곽이 윤부보다 작고보다 예리한 경우에 바람직하다. 또한, 그 크기로 인해, 눈꺼풀에 의해 가려지는 경향이 적습니다. 적외선 영상 기술의 주된 단점은 주변 적외선 조명으로 인해 낮에는 야외에서 사용할 수 없다는 것입니다. 이 논문에서는 알고리즘 개발을 적외선 스펙트럼 이미징 기법에 중점을 두지 만 이러한 기술을 가시 스펙트럼 이미징으로 확장하는 것을 목표로합니다.

적외선 시선 추적은 일반적으로 밝은 동공 또는 어두운 동공 기술을 사용합니다 (그러나 밝은 동공 기술과 암 동공 기술의 결합 된 사용 참조). 밝은 동공 기술은 카메라 축 위에 있거나 카메라의 축 근처에있는 광원으로 눈을 조명합니다. 이러한 조명의 결과는 동공이 눈의 광 반사 특성으로 인해 밝은 영역으로 명확하게 구분된다는 것입니다. 어둠 - 동공 기술은 동공이 이미지의 가장 어두운 영역이되도록 축을 벗어난 소스로 눈을 조명하지만, 공막, 홍채 및 눈꺼풀은 모두 상대적으로 더 많은 조명을 반사합니다. 두 방법 모두에서, 각막의 오프 조명 원의 제 1 표면 정반사 (아우터-눈의 가장 광학 소자)도 볼 수있다. 동공 중심과 각막 반사 사이의이 벡터는 일반적으로 동공 중심보다 오히려 부수적 인 척도로 사용됩니다. 이는 벡터 차이가 카메라와 소스가 동시에 움직이는 헤드 기어의 미끄러짐에 둔감하기 때문입니다 (아래의 검증 연구 결과 참조). 이 논문에서는 어두운 - 동공 기술을위한 알고리즘 개발에 초점을 맞추지 만, 우리의 알고리즘은 밝은 동공 기술에 쉽게 적용될 수있다.

**3. 관련 연구** 아이 추적 알고리즘은 피쳐 기반 (feature-based)과 모델 기반 (model-based) 두 가지 접근 방식으로 분류 할 수있다. 피쳐 기반 접근법은 눈 위치와 관련된 이미지 피쳐를 감지하고 지역화합니다. 지형지 물 기반 접근법의 공통성은 지형지 물의 존재 여부를 결정하기위한 기준 (예 : 임계 값)이 필요하다는 것입니다. 적절한 임계 값의 결정은 일반적으로 사용자가 조정하는 자유 매개 변수로 남겨 둡니다. 추적 된 기능은 알고리즘에 따라 크게 다르지만 대부분의 경우 강도 수준이나 강도 기울기에 의존합니다. 예를 들어, 어두운 - 동공 기술로 생성 된 적외선 이미지에서, 적절하게 설정된 강도 임계치는 동공에 대응하는 영역을 추출하는데 사용될 수있다. 동공 중심은이 확인 된 지역의 기하학적 중심으로 간주 될 수 있습니다. Intensity gradient는 가시 스펙트럼 영상에서 윤부를 검출하는데 사용되거나 [에서 동공 윤곽을 검출하는데 사용될 수있다

2]

적외선 스펙트럼 영상[12]. 그런 다음 타원을 이러한 특징 점에 적용 할 수 있습니다.

반면에 모델 기반 접근법은 기능을 명시 적으로 탐지하지 않고 오히려 이미지와 일치하는 가장 적합한 모델을 찾습니다. 예를 들어, 적분 차감 연산자를 사용하여 윤곽선과 동공 윤곽에 가장 잘 맞는 원형 [3] 또는 타원 [11]을 찾을 수 있습니다. 이 접근법은 원 또는 타원의 윤곽을 따라 미분의 적분을 최대화하는 모델 매개 변수 공간의 반복 검색을 필요로합니다. 모델 기반 접근법은 피처 정의 기준이 이미지 데이터에 적용되지 않는다면 피처 기반 접근보다 학생 중심의보다 정확한 추정을 제공 할 수 있습니다. 그러나,이 접근법은 국소 최소치를 가질 수있는 복잡한 파라미터 공간을 탐색 할 필요가있다. 따라서 모델 매개 변수에 대한 좋은 초기 추측 없이는 그라디언트 기법을 사용할 수 없습니다. 따라서, 모델 기반 접근법의 정확성에 대한 이득은 계산 속도 및 유연성 측면에서 상당한 비용으로 얻어진다. 그러나 모델 기반 접근법과 결합 된 다중 스케일 영상 처리 방법의 사용은 실시간 성능을 보장한다 [5]. **4. Starburst Algorithm** 이 섹션에서는 어두운 - 동공 적외선 조명에 대한 런타임 성능과 정확성 사이의 절충점을 얻기 위해 피쳐 기반 접근법과 모델 기반 접근법을 결합한 아이 추적 알고리즘을 제시합니다. 이 알고리즘의 목표는 동공 중심의 위치와 각막 반사를 추출하여 이러한 측정 간의 벡터 차이를 장면 이미지의 좌표와 관련시키는 것입니다. 알고리즘은 이미지에서 각막 반사를 찾아 제거함으로써 시작됩니다. 그런 다음 동공 가장자리 점은 반복적 인 피쳐 기반 기법을 사용하여 배치됩니다. 타원은 RANSAC (Random Sample Consensus) 패러다임 [4]을 사용하여 감지 된 가장자리 점의 하위 집합에 맞춰집니다. 그런 다음이 피쳐 기반 접근법에서 가장 적합한 피팅 매개 변수를 사용하여 이미지 데이터에 대한 적합성을 최대화하는 타원 매개 변수에 대한 로컬 모델 기반 검색을 초기화합니다.

**4.1. 노이즈 감소** 섹션 2에서 설명한 저가의 헤드 장착 아이 추적 장치를 사용하기 때문에 이미지에 나타나는 노이즈를 줄여야합니다. 노이즈, 샷 노이즈 및 라인 노이즈의 두 가지 유형이 있습니다.5적용하여 샷 노이즈를 줄*×* 표준 편차가 2 픽셀 인5 가우스 필터를입니다. 라인 노이즈는 가짜이며 정규화 계수를 라인별로 적용하여 라인의 평균 강도를 이전 프레임에서 얻은 실행 평균으로 옮길 수 있습니다.대한이인자 *C* 각 라인 *(L)에* 프레임의 *i는* 인

*C*(I,l)= *β* ̄*I*(I,l)+이 *-(β*1)*C*(I*-* 1,L)를(1) 여기서  ̄*I*(I,*L*)는 평균 선 강도이고 *β* = 0*이다.*이 잡음 감소 기술은 선택적이며수 있음에 유의하십시오

알고리즘이 잡음이 적은 영상을 캡쳐 할 수있는 아이 트랙커와 함께 사용될 때 제거 될.

**4.2. 각막 반사 탐지, 위치 파악**

**및 제거**

각막 반사는 눈 이미지에서 가장 밝은 영역 중 하나에 해당합니다. 따라서 각막 반사는 thresholding을 통해 얻을 수 있습니다. 그러나 관측자와 관측자간에 일정한 임계 값은 최적이 아닙니다. 그러므로 우리는 각 프레임에서 적응 형 thresholding 기술을 사용하여 각막 반사를 지역화합니다. 각막이 윤부에 거의 확장되기 때문에 각막 반사에 대한 검색을의 반 너비 인 정사각형 영역으로 제한 할 수 있습니다 *h* = 150 픽셀(매개 변수 값 관련 논의 섹션 참조). 시작하기 위해, 최대 임계 값은이 임계 값 이상의 값만 각막 반사 후보로 취하는 이진 이미지를 생성하는 데 사용됩니다. 가장 큰 후보 영역은 각막 반사에 기인 할 가능성이 있습니다. 그 이유는 다른 반사 효과가 눈꺼풀이 만나는 이미지 구석 근처뿐만 아니라 각막을 벗어나 매우 작습니다. 임계 값이 낮아짐에 따라 가장 큰 후보 영역과 다른 영역의 평균 영역 사이의 비율이 계산됩니다. 처음에는 각막 반사가 다른 영역보다 빠르게 커질 것이기 때문에 비율이 증가합니다. 각막 반사의 강도는 가장자리를 향해 단조롭게 감소하여이 성장을 설명합니다. 임계치가 낮을수록 일반적으로 잘못된 후보자가 증가합니다. 잘못된 후보가 더 두드러지고 각막 반사 영역의 크기가 커짐에 따라 비율이 떨어지기 시작할 것입니다. 가장 높은 비율을 생성하는 임계 값을 최적으로 취합니다. 각막 반사의 위치의 기하학적 중심 (의해 주어진다*xc, yc*이미지에서 가장 큰 영역)에 는 적응 적으로 결정된 임계 값을 사용하여.

크기가 작 으면 각막 반사는 이미지의 대략 원입니다. 국소화 단계에서 임계 값 영역을 사용하여 각막 반사의 대략적인 크기를 유도 할 수 있지만이 영역에는 일반적으로 각막 반사의 전체 프로파일이 포함되지 않습니다. 각막 반사의 전체 범위를 결정하기 위해, 우리는 각막 반사의 강도 프로파일이 2 변수 Gaussian 분포를 따른다고 가정합니다.반경 *r* 강도의 평균 감소가 최대 인을 찾고 가우스 (Gaussian)의 반경이 최대 반경 인 반경 (즉, 표준 편차 1의 반경)과 관련 시키면 각막 반사의 최대 범위를 2로 취할 수 있습니다*.*5는*R* 각막 반사 프로파일의 99 %를 캡처한다. 우리는찾을 수 *연구를* 그라데이션 괜찮은 통해

최소화검색을 ∫ *I(R* + *δ, XC,Y의C,*θ)*dθ* ∫ *I*(R*- δ, XC,Y의C,*θ)*dθ* (2)

*δ* = 1, 및 *I는*(R은,*X는, Y는,*θ)각도 픽셀 *θ에서의* 매개 변수에 의해 정의되는 원의 형상에 *r에있는*강도이며, *X*

3

1 **입력**: 제거 각막 반사추측으로아이 이미지

동공 중심 2 **출력의**: 특징점 집합 3 **절차 :** 4 반복 5 *단계 1 :* 6 시작점으로부터 연장 된 광선 따라 가기 7 각 지점에서 강도 미분 계산 8 파생물 *>* 임계 값이면 9 특징 지점 배치 10 광선을 따라 행진 중지 11 *스테이지 2 :* 12 1 단계에서 감지 된 특징점 3 월 13 일 시작 지점으로 되돌아 오는 광선을 따라 14 각 지점에서 강도 미분 계산 15 미분 *>* 임계 값이면 16 특징점 17 배치 광선에 따라 행진 중지 18 시작점 = 특징 지점의 기하학적 중심 19 시작점까지 수렴

그림 2 : 특징점 검출 방법 과 *y*. 검색은로 초기화됩니다 *r* = √*area / pi*. 여기서 *area* 는 임계 값 영역의 픽셀 수입니다. 검색은 빠르게 수렴합니다.

그런 다음 방사형 보간을 사용하여 각막 반사를 제거합니다. 먼저, 식별 된 각막 반사 영역의 중심 화소는 영역의 윤곽을 따른 강도의 평균으로 설정된다. 그런 다음 중심과 윤곽 사이의 각 픽셀에 대해 픽셀 강도가 선형 보간법을 통해 결정됩니다. 이 프로세스의 예가 그림 5에 나와 있습니다 (a와 b 비교 참조).

**4.3. 동공 윤곽 검출동공 윤곽을 검출** 우리는하는 새로운 피쳐 기반 방법을 개발했습니다. 알고리즘을 설명하는 의사 코드는 그림 2에 나와 있습니다. 다른 피쳐 기반 접근법은 전체 눈 이미지 또는 추정 된 동공 위치 주변의 관심 영역에 가장자리 감지를 적용하지만, 이러한 접근법은 학생 윤곽으로 계산적으로 낭비 적이 될 수 있습니다 이미지를 거의 차지하지 않습니다. 대신 우리는 동공 중심의 가장 좋은 추측으로부터 연장되는 제한된 수의 광선을 따라 모서리를 탐지합니다. 이 광선은 그림 3a에서 볼 수 있습니다. 이 방법은 어두운 동공 기술을 사용하여 적외선 조명으로 촬영 한 이미지에 존재하는 동공 윤곽의 고 대비 타원형 프로파일을 이용합니다.

각 프레임에 대해, 프레임에서 동공 중심의 가장 좋은 추측을 나타내는 위치가 선택됩니다. 첫 번째 프레임의 경우 수동으로 결정하거나 이미지의 중심으로 사용할 수 있습니다. 후속 프레임들에 대해, 이전 프레임으로부터의 동공 중심의 위치가 사용된다. 다음으로,따르는 미분 Δ *N* 이 시작점으로부터 반경 방향으로 연장되는= 18 광선을는 임계 값될 때까지 픽셀 단위로 개별적으로 평가됩니다 *φ* = 20이 초과. 우리가 어두운 동공 기술을 사용하고 있다는 것을 감안할 때, 긍정적 인 deriva

(b) (c) 그림 3 : 기능 탐지. (a) 동공 윤곽선 후보들 은 동공 중심의 최선의 추측으로부터 연장되는 일련의 광선의 길이를 따라 검출된다. 동공 컨투어 후보는 십자 기호로 표시됩니다. 두 개의 윤곽선 후보가 잘못되어 한 광선이 경계에 도달하여 후보를 생성하지 않습니다. (b) 각각의 동공 윤곽 후보에 대하여, 동공 윤곽 후보가 아닌 제 2 세트의 동공 윤곽 후보를 생성하는 광선 세트가 생성된다 (c) 동공 윤곽 상에 있지 않은 동공 윤곽 후보는 윤곽선 상에없는 추가 피쳐 포인트를 유도 할 수 있지만, 이들은 전형적으로 어떤 타원과도 일치합니다.

tives (광선이 확장됨에 따라 강도가 증가 함)가 고려됩니다. 이 임계 값을 초과하면 해당 위치에 기능 점이 정의되고 광선을 따라 처리가 중단됩니다. 광선이 이미지의 경계로 확장되면 특징점이 정의되지 않습니다. 후보 피쳐 포인트의 예시적인 세트가도 3a에 도시되어있다.

각 후보 특징점에 대하여, 전술 한 특징 검출 과정이 반복된다. 그러나 광선은로제한됩니다 *γ* = *±*원래 특징 점을 생성 한 광선 주위50 도로. 이러한 방식으로 복귀 광선을 제한하려는 동기는 후보 특징점이 실제로 (도 3b에 도시 된 바와 같이) 동공 윤곽 상에있는 경우, 복귀하는 광선은 동공의 반대편에 부가적인 특징점을 생성하여 모두 하나의 타원 (즉, 동공 윤곽)과 일치합니다. 그러나, 후보자가 학생이 아닌 경우 (예를 들어,도 3c 참조),이 프로세스는 임의의 단일 타원과 반드시 ​​일치하지는 않는 후보 후보 특징 점을 생성 할 것이다. 따라서,이 절차는 동공 윤곽선 상에없는 특징점들의 수에 대한 동공 윤곽상의 특징점들의 수의 비율을 증가시키는 경향이있다. 큰 Δ로 정의 된 피쳐 포인트가 동공 윤곽에 위치 할 확률이 높을 경우 (가장 강한 이미지 윤곽이므로) 반환되는 광선 수는 5설정됩니다*φ /*Δ로. 정의에 따라 특징점은 Δ의해 결정되기 때문에 최소 광선 수는 5 *>*= *φ에*입니다.

2 단계 기능 탐지 프로세스는 시작 지점에 대한 초기 추측에 대한 방법의 견고성을 향상시킵니다. 이것은 눈이움직임이때 문제가됩니다

4일

(a)

프레임에서 프레임으로의 위치를 ​​빠르게 바꿀 수 있기 때문에 눈. 이는 특히 낮은 프레임 속도에서 얻은 이미지에 적용됩니다. 예를 들어, 그림 4a에 표시된 것은 그러한 경우입니다. 초기 광선 세트는 동공 윤곽선상의 3 개의 특징점만을 검출하지만,이 3 점으로부터의 복귀 광선은 윤곽선상의 더 많은 점을 검출합니다 (그림 4b 참조). 특징점들의 결합 된 세트가도 4d에 도시되어 있으며, 윤곽선상의 점들의 개수는 윤곽선의 점들의 수를 초과한다. 그러나, 특징점은 초기화 점에 가장 가까운 동공 윤곽의 측면에 바이어스된다. 레이 프로세스의 또 다른 반복은 이러한 편향을 최소화 할 수 있지만 계산 부담은 각 반복마다 기하 급수적으로 증가하므로 비효율적 인 전략이됩니다.

이 시점에서 타원이 후보 지점에 맞추어 질 수 있지만 편향은 피팅에 심각한 오류를 유발합니다. 이러한 편향을 없애기 위해, 전술 한 2 단계 형상 검출 프로세스가 반복된다. 첫 번째 반복 이후의 각 반복에 대해 마지막 반복에서 모든 후보 특징 점의 평균 위치를 다음 시작 위치로 취합니다. 그림 4d의 빨간색 원은 두 번째 반복의 시작점을 나타냅니다. 두 번째 반복에 대해 감지 된 특징 위치가 그림 4e에 나와 있습니다. 강한 편향이 없음을 주목하십시오. 그림 4f는 중앙 위치가 실제 학생 중심으로 어떻게 빠르게 수렴되는지를 보여줍니다. 검출 된 특징점의 중심이보다 작아지면 반복이 중지됩니다 *d* = 10 픽셀. 예를 들어 대다수의 프레임을 차지하는 안구 고정 (eye fixation)과 같은 초기 추정이 학생 중심의 양호한 추정 일 때, 단지 하나의 반복 만이 요구된다. 초기 추정치가 좋지 않은 경우 일반적으로단지 몇 번의 반복 (*<* 수렴 (convergence)을 위해5)이 필요합니다.내에서 수렴에 도달하지 않으면 *i* 동공이 보이지 않을 때 깜박이는 동안 때때로 발생하는= 10 회 반복알고리즘이 중단되어 다음 프레임 처리를 시작합니다.

**4.4. 타원 피팅** 후보 피쳐 포인트 세트가 주어지면 알고리즘의 다음 단계는 가장 적합한 피팅 타원을 찾는 것입니다. 다른 알고리즘은 일반적으로 모든 특징 점 (예 : [20] 참조)에 대한 타원의 최소 자승 피팅을 사용하지만 피쳐 검출 단계에서의 총 오류는 결과의 정확성에 크게 영향을 줄 수 있습니다. 도 5c에 도시 된 검출 된 특징 점 및도 5d에 도시 된 최소 자법을 사용하여 결과적인 최적의 타원을 고려한다. 학생 윤곽선에없는 몇 가지 특징점은 피팅의 품질을 허용 할 수없는 수준으로 대폭 감소시킵니다.

이 문제를 해결하기 위해 모델 피팅 [4]에 RANSAC (Random Sample Consensus) 패러다임을 적용합니다. 우리의 지식에 따르면, 우리는 안구 추적의 관점에서 RANSAC을 처음으로 응용 한 것이지만, RANSAC은 다른 컴퓨터 비전 문제에 자주 적용됩니다 (예 : [7] 참조). RANSAC는에서 큰 값이지만 알려지지 않은 비율의 외래 값이있을 때 모델 피팅에 효과적인 기법입니다

(e) (f) 그림 4 : .피쳐 탐지. (a) 원래의 시작점 (노란 원)은 광선 (파란색)을 쏘아 후보 학생 점 (녹색 십자)을 생성합니다. (b & c) 후보 학생 포인트는 더 많은 후보 학생 포인트를 탐지하기 위해 시작점을 향해 광선을 다시 쏘습니다. (d) 모든 후보 학생 포인트가 표시됩니다. 이 위치의 평균은 빨간색 원으로 표시됩니다. 이 위치는 다음 반복을 시드합니다. (e) 두 번째 반복의 결과. (f) 모든 반복의 시작 위치는 빠른 수렴을 보여준다.

측정 샘플. inlier는 모델링 된 메커니즘에 기인 한 데이터의 샘플이며 outlier는 오류를 통해 생성 된 샘플이며 고려 대상이 아닌 다른 메커니즘에 기인합니다. 우리의 응용에서, inliers는 동공 윤곽에 해당하는 검출 된 모든 특징 점이며 outlier는 눈 뚜껑과 눈 사이의 윤곽과 같은 다른 윤곽에 해당하는 특징 점입니다. 최소 제곱 법은 모든 샘플이 인라이어이고 모든 에러가 측정 에러에 기인한다고 가정하기 때문에 사용 가능한 모든 데이터를 모델에 적합하게 사용합니다. 반면 RANSAC는 이상 치의 가능성을 인정하고 모델의 적합성을 위해 데이터의 하위 집합만을 사용합니다. 자세히 말하자면, RANSAC은 데이터의 작지만 임의적 인 부분 집합을 많이 선택하고, 각 부분 집합을 사용하여 모델에 적합하게하며, 전체 데이터 집합과 가장 일치하는 모델을 찾습니다. 이 모델과 일치하는 데이터의 하위 집합은 합의 집합입니다.

경우에 따라, 우리의 2 단계 특성 검출 과정은 극히 적은 이상치 (예를 들어, 그림 5e 참조)를 가져오고, 다른 경우에는 이상치가 훨씬 더 많습니다 (그림 5f 참조). 그러므로 우리는 동공 윤곽에 가장 잘 맞는 타원을 찾기 위해 RANSAC 패러다임을 사용하는 것이 중요합니다. 다음 절차는반복 *R* 번됩니다. 먼저,5 개의

5

(a) (b)

(c) (d)

타원의 모든 매개 변수를 결정하는 데 필요한 최소 샘플 크기 인 경우샘플이 탐지 된 피쳐 세트에서 무작위로 선택됩니다. 정규화 된 특징점 좌표 [7]로 생성 된 원추형 구속 조건 행렬에 대한 SVD (Singular Value Decomposition)는이 다섯 점에 완벽하게 맞는 타원의 매개 변수를 찾는 데 사용됩니다.

타원의 매개 변수가 허수이면 타원 중심이 이미지 외부에 있거나 장축이 보조 축의 두 배보다 크면 더 이상 해당하지 않을 때까지 5 개의 다른 점을 무작위로 선택합니다. 그런 다음이 모델과 일치하는 데이터 집합의 후보 피쳐 포인트 (즉, 인위 수)의 개수가 계산됩니다. Inliers는 타원에 대한 대수 거리가 일부 임계 값보다 작은 샘플 점입니다 *t*. 이 임계 값은 우리의 지형지 물 탐지기의 특성을 기반으로 예상되는 오류의 확률 적 모델에서 파생됩니다. 우리의 특징 검출기의 평균 오차 분산은 약 1 픽셀이고이 오차는 평균이 0 인 가우시안으로 분포되어 있다고 가정합니다. 따라서 샘플이 정확하게 inlier로 분류되는 95 % 확률을 얻으려면 문턱 값은를 갖는로부터 도출되어야한다 *χ*2 1 자유도분포. 임계 값 거리가 1이*됩니다.*98 픽셀.후에 *R* 반복, 가장 큰 합의 집합을 가진 모형이 사용된다. 모든 가능한 특징점 조합을 평가하는 것은 종종 계산적으로 실행 불가능하기 때문에, 무작위로 선택된 부분 집합 중 적어도 하나가 오직 유일한 값을 포함한다는 것을 보장하도록 시도 할 임의의 부분 집합의 수를 결정해야한다. 이는 확률으로 보장 할 수 있습니다 *p* = 0*.*99,

*R* = log (1 *- p*)

log (1 *- w*5) (3)

여기서 *w* 는 표본에서의 inliers 비율이다.않았지만 *w* 는알려지지 *선험적으로*, 그 하한은 반복에서 임의의 모델에 대해 발견 된 최대 inlier 수에 의해 주어지며, 따라서 *R* 은 반복이 진행됨에 따라 식 3을 사용하여 초기에 매우 크게 설정되고 낮출 수있다. 필요한 반복 회수 후에, 타원은 가장 큰 합의 집합에 맞는다 (예, 그림 5g 참조).

**4.5. 모델 기반 최적화** RANSAC 적합성의 정확성이 많은 눈 추적 애플리케이션에 충분할 수 있지만 타원형 피팅 결과는 피쳐 감지에 의존하지 않는 모델 기반 최적화를 통해 향상 될 수 있습니다.의 매개 변수를 찾기 위해 *a, b, x, y, α* 최적 피팅 타원우리는

*-*∫ *I*(*a* + *δ, b* + *δ, α, x, y, θ*)*dθ* ∫ *I*(*a - - δ, α, X, Y,*θ)*dθ* (4)

여기서, *δ* = 1, *I*(A,*B, α, X, Y는,*θ)에각도 픽셀 강도 *θ에서의* 의해 정의되는 타원 형상에매개 변수 *a*,*b*,*x*,*y* 및 *α*. 검색은 RANSAC 적합도에 따라 결정되는 가장 적합한 맞춤 타원 매개 변수로 초기화됩니다.

모델 기반 학생 중심 1 차 2 차 3 차

좁은 FOV 0.507 6.572 10.362 와이드 FOV 0.591 7.527 12.316 모델 기반 벡터 차이 1 차 2 차 3 차

좁은 FOV 0.471 0.981 1.204 와이드 FOV 0.515 1.203 1.565

표 1 : 검증 연구의 정확도 결과

**4.6. Homo graphic mapping and calibration** 씬 이미지에서 사용자의 시점을 계산하기 위해, 씬 이미지의 위치와 아이 위치 측정 값 (예 : 동공 중심과 각막 반사 간의 벡터 차이)을 매핑해야합니다 결정되어야한다. 안구 추적 방법의 전형적인 절차는 보정 절차 [17]를 통해이 관계를 측정하는 것입니다. 캘리브레이션 동안, 사용자는 씬 이미지 내의 위치들이 알려지는 다수의 씬 포인트들을 볼 필요가있다. 사용자가 각 장면 점고정하는 동안 *s* = (*xs, ys,*1)을눈 위치 *e* = (*xe, ye,*1) 가 측정됩니다 (동종 좌표를 기록하십시오). 선형 호모 그래픽 매핑을 사용하여 두 세트의 점 사이의 매핑을 생성합니다. 이 매핑 *H* 는 3 *×* 3 행렬이며 8 자유도를 갖습니다.의 엔트리를 결정하기 위해, *H*제약 매트릭스는 측정 된 포인트 대응을 이용하여 생성된다. 각각의 서신은 두 개의 구속 조건을 생성하므로 4 개의 대응 성이를 풀기에 충분하다 *H* 에 대한 척도[7]. 제약 행렬의 영 공간은 SVD를 통해 결정될 수 있으며,제공한다 *H를*. 일단이 맵핑이 결정되면, 임의의 프레임에 대한 장면에서 사용자의 시점을로 설정할 수있다 *s* = *H e*.을 사용합니다 *3x3* 안구 운동의 정확한 예측을 보장하기 위해 장면 이미지에 균일하게 분포 된 눈금 교정 점. 이 경우 미지수보다 많은 제약이 있으며 SVD는매핑생성합니다 *H* 대수적 오차 거리를 최소화하는를.

**5. 알고리즘 검증 알고리즘**

의 성능을 검증하기 위해 안구 추적 평가가 수행되었습니다. 비디오는 2 장에서 설명한 머리 장착형 아이 트랙커에서 기록한 반면, 세 명의 저자는 랩톱 컴퓨터에 두 개의 영화 예고편을 보았습니다. 각 트레일러를 시청하기 전과 후에 사용자는 턱 받침대에 머리를 올려 놓고 약 60cm 떨어진 위치에있는 흰색 보드에 일련의 9 개의 보정 표시를 고정했습니다. 평가는 각 사용자에 대해 두 번 수행되었습니다.

두 번째 평가좁은 시야 장 (56*o* 에서 장면 카메라에 사용 된FOV)을 넓은 시야 렌즈 (111로 교체하여*o* FOV 및 중요한 방사 왜곡)렌즈의 비선형 왜곡에 기인 한 아이 추적 품질. 평가 중에 캡처 된 비디오는 http://hcvl.hci.iastate.edu/openEyes에서 볼 수 있습니다.

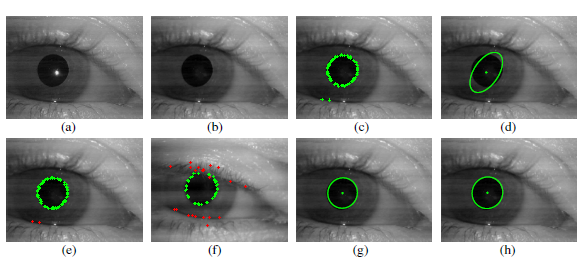
그림 5 : (a) 노이즈 감소가있는 원본 이미지. (f) (b) 각막 (g) (h) 반사가 제거 된 이미지. (c) 후보 피쳐 포인트. (d) 최소 제곱 법을 사용하여 타원형으로 고정. (e) RANSAC에 의해 구분 된 inliers (십자형 녹색) 및 outliers (적색 십자형). (f) 더 많은 이상치가있는 또 다른 예. (g) 유일 타원 (inliers)만을 사용하는 타원형. (h) 모델 기반 최소화를 사용하여 가장 적합한 타원. 설명을 위해이 그림에서 *N* = 12를 사용합니다. 

표 1에 표시된 것은 교정 그리드의 1 차, 2 차 및 3 차 관찰에서 얻은 정확도 추정치입니다. 정확도는 9 개의 모든 교정 점에 대해 평균 된 장면 이미지의 보정 표식의 실제 위치와 추정 된 시선 점 사이의 거리로 측정됩니다. 유효성 검사 결과는 동공 센터를 사용할 때와 동공 중심과 각막 반사 중심 간의 벡터 차이를 사용하여 시선을 예측할 때 별도로 표시됩니다. 모눈의 첫 번째보기는 모든 예측에 대한 호모 그래피 매핑을 생성하는 데 사용됩니다.

주목할 첫 번째 결과는 동공 중심 대신 각 동 중심과 각막 반사 중심 간의 벡터 차이가 사용될 때 오류가 현저히 낮다는 것입니다. 오류는 동공 중심만을 기준으로 한 시선 견적 점에 대한 각 보정 후에 증가하는 경향이 있습니다. 이것은 놀라운 것이 아니며 헤드 기어의 미끄러짐 때문입니다. 두 번째로 주목해야 할 점은 오차 패턴이 넓은 시야와 좁은 시야의 렌즈 모두에서 비슷하다는 것입니다. 비선형 성이 중요하지 않다는 것을 암시합니다. 필요한 경우 표준 이미지 처리 기술을 사용하여 각 프레임의 방사형 왜곡을 제거하여이 오류를 수정할 수 있습니다.

**6. 토론** 우리는 특징 기반과 모델 기반 접근을 결합한 안구 추적을위한 하이브리드 알고리즘을 개발했다. 각막 반사와 동공은 적응 형 피쳐 기반 기법을 통해 위치합니다. 그런 다음 RANSAC 패러다임을 적용하여 거친 피처 검출 오류가있는 경우 타원 피팅의 정확성을 극대화합니다. 마지막으로, 모델 기반 접근법이 적합성을 더욱 구체화하기 위해 적용됩니다. 우리는 알고리즘이 저가의 머리 장착 안구 추적기에서 얻은 비디오에서 잘 작동 함을 나타내는 검증 연구를 실시했습니다.

우리는 여전히 무료 매개 변수의 변형에 대한 알고리즘의 견고성을 탐구하는 과정에 있습니다. 예를 들어, 검증 연구에 사용 된 지형지 물 감지 임계 값은 손으로 결정되었지만 상대적으로 강건 해 보였으므로 다른 사용자에 대해 다르게 설정하지 않아도됩니다. 그러나이 임계 값은 주어진 eye tracker에 대해 조정해야 할 가능성이 높습니다. 기능 탐지 프로세스에서 다시 광선 수를 조작하는 효과를 조사하지 않았으므로 광선을 많이 사용하면 런타임 성능에 대한 비용을 최소화하면서 견고성을 높일 수 있습니다. 우리는 현재 시선의 예상 지점의 품질에 이러한 매개 변수의 영향을 조사하고 있습니다.

현재의 구현에서 많은 개선이 이루어질 수 있습니다. 예를 들어, 이미지에서 각막 반사를 제거하는 대신 꽤 많은 시간이 소요될 수 있으므로 알고리즘의 다른 단계에서 각막 반사 영역을 무시할 수 있습니다. 현재 알고리즘이 본질적으로 독립적으로 이미지를 처리한다는 점을 감안하면 추가 개선의 여지가 있습니다 (단, 이전 프레임에서 동공 중심의 추정이 현재 프레임의 동공 중심을 가장 잘 추측하는 데 사용됨). 예를 들어, 우리는 칼만 필터를 사용하여 학생 중심의 예측을 통해 얻을 수있는 향상을 모색하고 있습니다. 그러나 낮은 프레임 속도와 높은 안구 움직임 속도를 감안할 때 하드웨어에 대한이 기술의 잠재적 이점을 예측하기는 어렵습니다. 우리는 또한 자동 교정을 연구 중입니다. 현재 캘리브레이션에는 장면 이미지에 캘리브레이션 포인트의 위치를 ​​표시하는 수동 입력이 필요하며 이는 귀찮은 작업이 될 수 있습니다. 이미지 처리 기술을 사용하여 장면 이미지의 교정 교차점을 자동으로 감지 할 수 있습니다.

우리의 연구는 범용 하드있는 안정적인 안구 추적 알고리즘을 개발하는 것을 목표

7

(a) (b) (c) (d)

웨어에서 실행될 수 있고 일상적인 인간 - 컴퓨터 인터페이스에서 널리 사용될 수로합니다. 자유롭게 사용할 수있는 안구 추적 소프트웨어가 부족하다는 점을 감안할 때, 우리는 GNU 공개 라이선스 (GPL)에 따라 공개 소스 소프트웨어 패키지에서 알고리즘 알고리즘을 구현할 수 있도록하고 있습니다. 이 소프트웨어는 웹 사이트 http://hcvl.hci.iastate.edu/openEyes에서 다운로드 할 수 있습니다. 우리는 오픈 소스 안구 추적 소프트웨어와 기존의 구성 요소로 제작 된 저비용 안구 추적 시스템을 결합하여 인터페이스 설계자들이 안구 운동의 잠재력을 탐색 할 수있을 것이라고 기대합니다 이는 인터페이스를 향상시키고, 이는 차세대 휴먼 - 컴퓨터 인터페이스에서의 안구 추적에 대한 역할을 증가시킬 것이다.