

〔公 開〕

TR-C-0070

視点及び視線検出のための  
特徴点実時間抽出処理と  
高精度化の検討

伴 野 明  
Akira TOMONO

岸 野 文 郎  
Fumio KISHINO

1 9 9 1 . 1 1 . 1 2

A T R 通 信 シ ス テ ム 研 究 所

— 目 次 —

1. まえがき	1
2. 画像処理による視点・視線の検出	
1. 検出の基本方針	2
2. 入力ツールとしての目標特性	5
3. 特徴点の撮影方法と強調処理	
1. 瞳孔	8
2. 顔マーク	14
3. 角膜反射像	14
4. 近赤外 3 板式 CCD カメラの試作	
1. 照明系	17
2. プリズム光学系	17
3. 駆動系とカメラ特性	19
5. 中心位置計算の高精度化	
1. 楕円近似による輪郭データの補正方法	21
2. シミュレーション実験	23
6. むすび	26

# 視点及び視線検出のための 特徴点実時間抽出処理と高精度化の検討

伴野 明, 岸野 文郎,

## 概要

本資料では、視点の検出を透視投影表示に、また、視線の検出を表示画面の直接指示に利用する場合を想定して、これら検出に必要な眼球の特徴点を画像処理により抽出し、画像上の位置を実時間で抽出する手法、及び、特徴点の中心位置を高精度で計算する手法について述べる。特徴点としては、瞳孔、角膜反射像、顔マークを対象とし、近赤外照明を用いてこれらを撮影した。近赤外照明は、眼への生理的影響が懸念されるため、少ない光エネルギーで効率良く撮影する方法を検討した。瞳孔の抽出では、眼底反射光を捉える撮影方法を用いた。瞳孔像の輝度は波長が短い程高い、また、長い波長は輝度には、あまり寄与しないため遮断するのが望ましい。また、外部照明の影響を受けずに、瞳孔を安定に抽出するためには、瞳孔が明るい画像と暗い画像を同時に撮影し、差分処理する方法が有利であり、このために、波長が僅かに異なる照明光を利用した。角膜反射の抽出では、偏光照明を利用した。反射光の中から偏光成分を遮断した画像と通過させた画像を同時に撮影し、これらを差分処理することにより角膜反射像が強調され簡単に2値化できる。各々の特徴点の抽出に必要な差分処理、2値化処理、及び重心位置計算は、パイプライン画像処理装置を用いて行い、40ms以下で実現した。特徴点の中心位置検出精度を上げるため、楕円近似を適用し、シミュレーション実験を行って考察し、視点・視線検出に必要とされる精度が得られる見通しを得た。

## 1. ま え が き

筆者等は、将来の電気通信の一形態として、仮想的な会議の”場”（仮想空間）をコンピュータグラフィックスによって作成し、その中に居るかのような感覚で会議や協調作業を行うことができる臨場感通信会議システムの構想を提案している[1]。このシステムでは、仮想空間を臨場感が高く、操作性の良いものにすることが本質的に重要となる。これには、受け手主体の表示方法を検討する必要がある、視点及び視線検出技術は重要な位置を占める。例えば、観察者の視点や視線の検出結果を表示方法にフィードバックして、運動視、眼鏡なし立体視、広視野・高精細表示[2]などを実現できれば、仮想空間の知覚はより自然なものになる。また、人の眼は表示画面の中の興味ある対象を求めて絶えず動作しているので、視線の動きの分析から操作意図が推定できれば、議題となっている表示内容を自動的に更新するなど、柔軟な応答を返す操作性のよいインタフェース機能が可能となる。

視点や視線をこのような目的に使用するためには、機器を装着せずに、且つ頭部の動きを許容する検出法が必要であり、著者らは、画像処理による手法を検討している[3]。視点や視線をカメラ画像から求めるためには、検出に適した特徴点を安定、且つ高速に抽出し、画像上でその位置を正確に求めることが重要である。

ところで、上記のような通信システムのインタフェース環境では、室外光や室内照明さらに表示装置からの光の放射などがあるため人物の照明条件は様々に変化し、特徴点抽出に必要な画像が常に安定に撮影できるとは限らない。入力画像の質が悪ければ、雑音除去などの処理が複雑になり、高速化や高精度化を損なうことになる。

特徴点を効率良く抽出するためには、眼に知覚されない近赤外光を顔や眼に照射するなどして、制御された照明条件のもとで、撮影する方法が効果的である。しかし、近赤外光の眼への生理的影響については不明な点も多く一般的には強い照明は望ましくないと考えるべきである[4]。特に、本目的のように、視点や視線検出手段をインタフェースツールとして位置づけ、長時間、連続的に使用する場合には、できるだけ弱い照明で特徴点を抽出できることが望ましい。

また、眼の中の特徴点は瞼や睫毛などにより遮られ、完全な形で撮影されない、つまり部分的に欠落することが多いため、欠落を補って正確な位置を求めることも必要である。そこで、本論文では、視点や視線の検出を臨場感表示や、視線による直接指示に応用する場合を想定して、画像処理により、このための特徴点を効率良く抽出し、実時間でその画像上の位置を精度良く求める手法について述べる。2. では、視点・視線検出のアプローチ、目標とする検出精度と速度などについて述べる。3. では、特徴点として有望と考えられる瞳孔、角膜反射像、および顔マークの抽出法について述べる。4. では、近赤外照明で特徴点を効率良く撮影する3板式CCDカメラの構成について示し、背景に対して特徴点を強調する信号処理方法について述べる。また、特徴点の強調化処理とその中心位置の検出を実時間でを行うパイプライン画像処理装置の構成について示す。5. では、

撮影画面上で特徴点の中心位置を高精度で求めるための処理方法について述べる。

## 2. 画像処理による視点・視線の検出

### 2.1 検出の基本方針

#### (1) 特徴点位置計測システム

図1はインタフェースの環境で使用する特徴点位置計測システムの構想を示したものである。

特徴点を効率良く抽出するために照明装置を利用し撮影する。特徴点の撮影画面上の位置をもとに、ステレオ画像計測などによって、その空間位置をカメラ座標系で求める。次に、これらの位置を表示スクリーンを一意に記述できる座標系（スクリーン座標系と呼ぶ）に変換する。ここで、カメラ座標系とスクリーン座標系との関係は予め精度良く求めておくこととする。スクリーン座標系で求めた瞳孔の位置は視点として利用し、また、顔、瞳孔、角膜反射像、などの特徴点の位置は視線検出に利用する。以下、スクリーン座標系で求めた特徴点の位置を用いて、視線を検出する手法の例を簡単に示す。

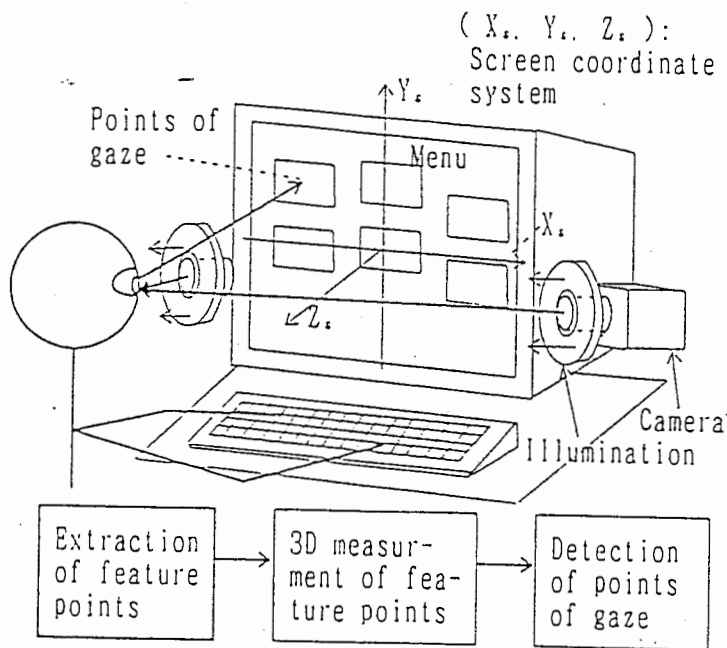


図1 非装着視線検出のアプローチ

Fig. 1 Approach to non-contact eye movement detection.

## (2) 顔と瞳孔を特徴点とする視線検出法

瞳孔は後述のように直接カメラ撮影できるので、画像計測でその位置を求めることができる。眼球中心は前述のように仮想的な点であり直接計測できないが、頭部の中で動かない点と考えると、顔の位置と方向から求めることができる[5]。顔上に3つの特徴点を設け、スクリーン座標系でこの空間位置を求めると、顔は、図2〔A〕のように、三角形でモデル化できる[6]。ここで、スクリーン座標系で求めた三角形の重心を顔の位置(図中:  $Fg(s)$ )、法線方向を顔の向きとすると、スクリーン座標系に一意に変換可能な顔座標系を作ることができる。眼球中心を顔座標系において固定ベクトル(眼球中心ベクトル  $E_c(f)$ )と呼ぶ)で表すと、眼球中心の位置はスクリーン座標系で求まるため、眼球中心と瞳孔中心( $P_u(s)$ )を結ぶ直線を光軸として求めることができる。光軸を近似的に視線とすることもできるが、視線と光軸との傾きを考慮する場合には、光軸ベクトルに所定の傾きを重畳し、視線ベクトルとすることもできる。

## (3) 瞳孔と角膜反射像を特徴点とする視線検出法

視線は、眼球回転中心の空間位置と眼の回転角が分かれば求まる。眼球回転中心は、瞳孔または角膜反射像の空間位置から近似的に求め、回転角は以下のように求める方法が考えられる。

光源と眼の距離を十分に離すと、眼に入射する光はほぼ平行光になるため、角膜反射像は、光源の方向から観察した場合、図2〔B〕に示すように、光源と角膜曲率中心  $K$  とを結ぶ線上で、 $K$  から  $r/2$  ( $r$ :角膜曲率半径)の点に生ずる。ここで、眼球中心  $E_c$  から瞳孔までの距離と、 $E_c$  から角膜曲率中心  $K$  までの距離を比較すると、前者の方が長いため、眼球の回転に対して、瞳孔は大きく、角膜反射像は小さく動く。従って、眼の中の2つの特徴点を拡大して撮影し、これらの相対的な動きを計測すれば眼球の回転角が求まる。

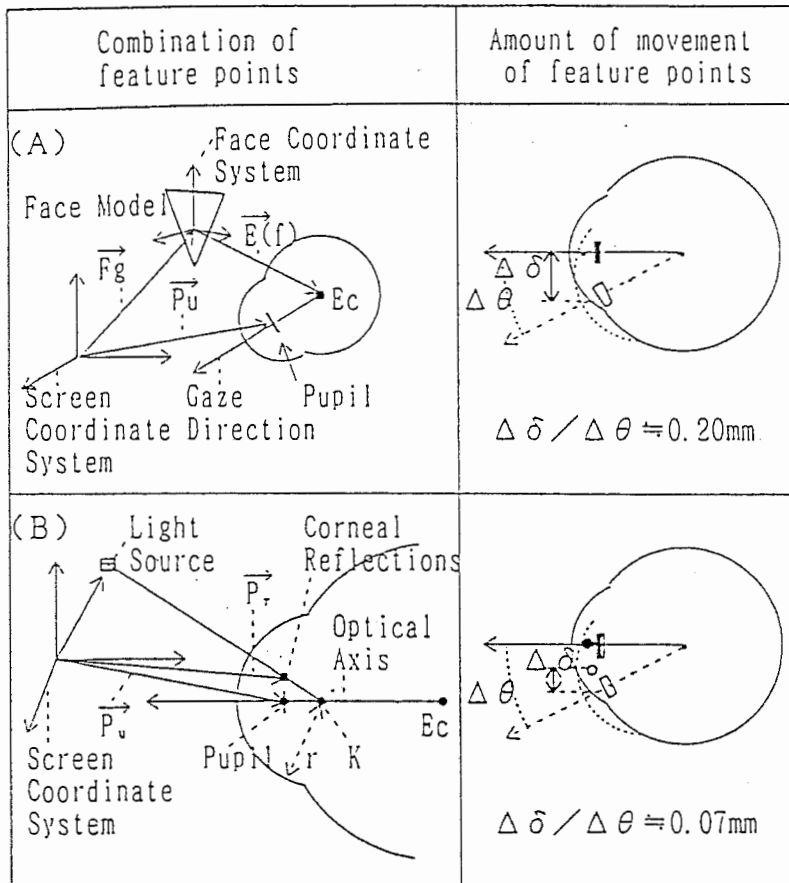


図2 2種類の視線検出法と特徴点の相対的な移動量

Fig.2 Eye movement detection method and amount of movement relative to the rotation of the eye.

## 2. 2 入力ツールとしての目標特性

視点位置情報の運動視CGへの利用[1]、視線情報の指示入力への利用を例として、必要とされる検出精度、検出速度について考察する。

### (1) 検出精度

視点： 運動視CGの実現では、3次元世界モデルを観察者の視点を基準にして透視投影する必要がある。瞳孔を視点と考え画像処理によりその位置を検出する場合、瞳孔抽出方法によっては誤差が生ずる。視点検出誤差が大きいと、透視投影画像はゆらいで観察されるため、不自然な感じを与える。そこで、視点の微小な変化と視覚特性の関係を実験により求めた。

図3に実験方法を示す[7]。著者らは、仮想空間に会議室を生成し、この中で、通信会議を行う構想を持っているため、3次元モデルとしては、会議室に相当する大きさの部屋に、注視対象として単純な形状の球がある世界を用い、70inchのスクリーンに透視投影した。視点1は、観察者の眼の位置で、スクリーンまでの距離は約1m、球までの距離は、約10mである。また、スクリーン上の球の大きさは視覚にして約5度である。従って、観察者は、球を中心視で捉え、部屋の壁を周辺視で捉えることになる。

実験では、観察者の頭部を視点1に眼球が位置するように固定し、観察者には球を注視するように指示した。この状態で、眼の計測位置に誤差がある場合をモデル化するため、視点1を30回/秒の速さで微小に変化させ(例：視点2)、それにともなって透視投影画像を更新して表示し、画像のゆれが知覚されるか否か評価した。視点は、視点1を中心にして、所定の振幅を設定し、この範囲内においてランダムに変化させた。評価は、全く知覚できない：スコア-4、少し知覚できる：3、少し気になるが許容できる：2、明らかに気になる：1、の4段階である。なお、被験者は、ランドルト視力が1.0以上の3名である。

図4に、視点の移動振幅に対する被験者の平均スコアを示す。同図から以下のことが分かる。

ア) 0.3mm以下の振幅では、画像のゆれは全く知覚できないが、2.0mm以上では明らかに気になる。

イ) ゆれを感じるが許容できるスコア-2に相当する視点の変化は約1mmである。

なお、本実験では、視点から表示対象までの距離が大きくなるに従って、同じ頭部の移動量に対する表示画面上の物体の動きは大きくなるため、ゆれを感じ易くなる。注視対象である球は、10mに置いたが、実際の会議室の奥行きは5m程度であり、この場合の評価は図4よりいくらか良くなると思われる。

以上の実験から、時系列で変化する視点の相対的な位置検出精度としては、1mm以下が望ましいと言える。なお、運動視CGにおいて、スクリーン座標系での絶対的な視点検出精度については、利用目的によっても異なるため省略する。

視線： 視線による指示入力では、ワープロ文字の数個程度、即ち画面上で半径1~2cmの範囲程度が指示できれば広い応用が考えられる。この範囲は、表示画面との視距離を60cmとすると1~2度に相当する。視線は、前述のように眼球の回転中心位置と回転角で決まるが、スクリーン座標系における特徴点位



置の検出誤差は、概ね眼球回転中心のずれに相当し、視線ベクトルの平行移動方向の誤差として作用する。一方、特徴点間の相対的な位置（例えば、瞳孔位置を基準にした場合の角膜反射像の位置）の誤差は、概ね視線の方位角誤差として作用する。

スクリーン座標系における特徴点位置の検出誤差については、便宜上、文字1個程度（約5mm）を当面の目標とする。また、特徴点間の相対的な位置検出誤差については、以下のような近似で見積もった[3]。

2. 1で述べた、2種類の視線検出法において、眼球の光軸上に照明を配置し、同じ方向から撮影する条件で、眼球を左右に回転したとき、回転角1度あたりの特徴点の移動量は、図2のようになる。なお、眼球構造パラメータは、文献8を参考に角膜頂点から眼球中心までの距離を13.3mm、瞳孔と眼球中心の距離を9.8mm、角膜曲率半径を7.7mmとした。従って、上記の眼球回転角の検出目標（1～2度）を踏まえると、特徴点間の相対位置は、顔の3点と瞳孔を特徴点とする視線検出法では、0.4mm以下、また、瞳孔と角膜反射像を特徴点とする視線検出法では、0.15mm以下で各々検出する必要がある。

## (2) 検出速度

運動視表示では、視点検出から表示までの時間が大きくなると違和感を生ずる。いくつかの評価画像について検討例があるが、自然物体を撮影したカメラ画像を用いた運動視表示実験では、遅れが知覚される検知限は約60msとされている[9]。

指示入力では、眼は跳躍運動を繰り返して目標を探索する。跳躍運動の性質として、1回の注視時間は200～350msが多く、1回跳躍が起きると、その後200msの不応期があることが知られている[10]。そこで、この時間を視線を検出してからシステムが応答するまでの目標時間と考える。

以上、視点・視線検出の用途に対する目標を表1にまとめて示す。高速で精度の高い検出が必要であり、特に、特徴点間の相対的な位置検出精度を上げるためには、撮影画像の中の特徴点を安定に抽出し、その位置を正確に求めることの重要性が分かる。以下、この目標を踏まえた特徴点の抽出処理について述べる。

表1 視点・視線検出の用途と目標特性

用途	特徴点の検出精度		検出速度
	相対位置	絶対位置	
運動視CGのための視点検出	<1.0 mm	—	< 60 ms
指示入力のための ①顔と瞳孔を用いた視線検出	<0.4mm	<5.0 mm	< 200ms
②瞳孔と角膜像を用いた視線検出	<0.15mm		

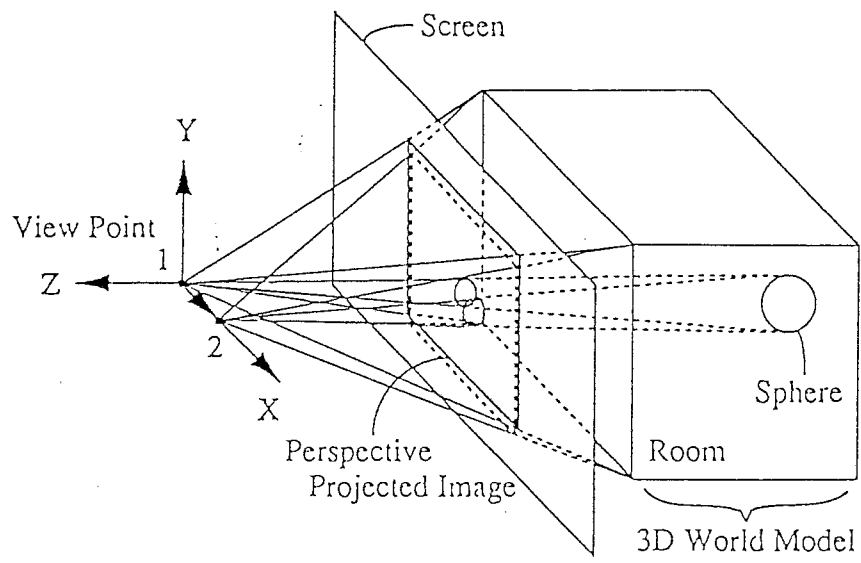


図3 実験方法  
Fig. 3 Experiment method

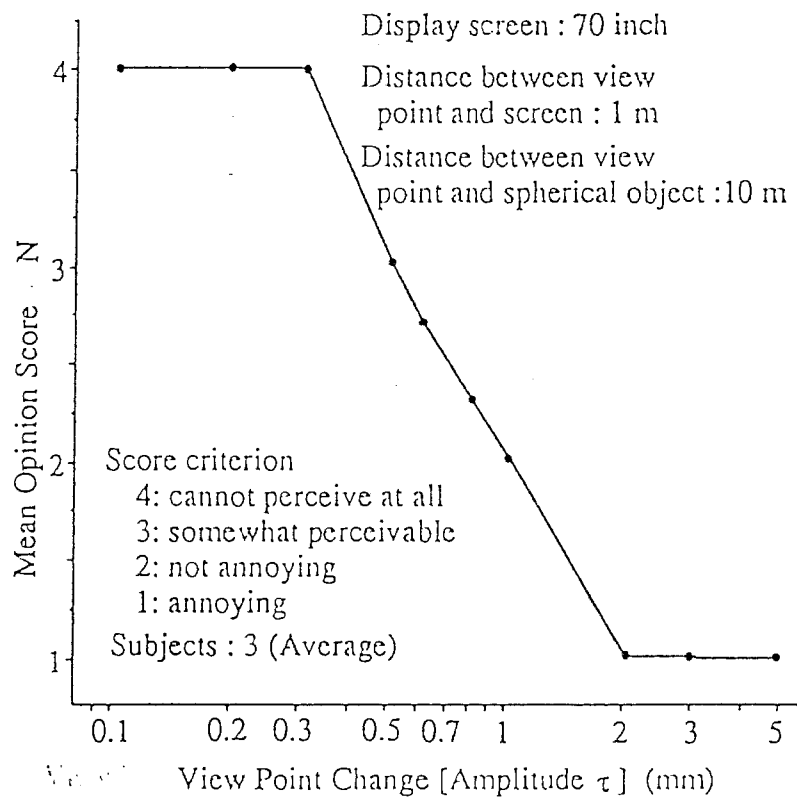


図4 視点の変化がゆらぎ知覚に及ぼす影響

Fig. 4 Evaluation of flicker sensation caused by small changes in viewpoint

### 3. 特徴点の撮影方法と強調処理

近赤外照明を用いて、瞳孔、顔をモデル化するための顔マーク、角膜反射像、の3つの特徴点を各々強調して撮影する手法について述べる。

#### 3.1 瞳孔

瞳孔を虹彩と区別して撮影するためには、照明装置を用いて、瞳孔から光を眼球内に入れ、眼底で反射して再び瞳孔から出る光をカメラで捉える方法が効果的である[11]。この場合、瞳孔は虹彩に対して明るく撮影される。

この方法では、カメラレンズの光軸に対する照明装置の位置が瞳孔像の輝度分布に大きく影響する。そこで、筆者らは先に、これらの関係を求めるシミュレーション手法を開発し、照明装置をレンズ有効径内に光軸に対して対象に配置することにより、眼球の大きな回転に対しても、瞳孔全体が均一に明るく撮影されることを示した[12]。しかし、このシミュレーションでは、眼底での光反射特性を一樣としており、照度や波長と眼底での反射特性の関係、また、カメラの感度特性などについては、考慮していない。このため、照明装置の最適な配置条件は分かっても、照度や波長と瞳孔像の明るさとの関係については、明らかではなかった。そこで、これらに着目して実験を行い、瞳孔を効率良く撮影するための条件を求めた。

##### (a) 照明強度と瞳孔像

図1に示したように、照明は眼の前方から行うため、知覚され難い近赤外光を用いる必要がある。しかし、近赤外光は眼に生理的影響を与える恐れがあるため、できるだけ弱い照明が望ましい。そこで、瞳孔撮影に必要な照明強度を図5の実験系を用いて調べた。照明の光源には、近赤外に比較的強いスペクトルを持ち、光強度を広い範囲で設定できるハロゲンランプを用いた。カメラレンズの光軸付近から照明するため、レンズの前にはハーフミラーを45度の角度に設け、光源は光軸に対して直角に設けた。また、光源とハーフミラーの間には、眼近傍に光を集めるためのレンズと、可視遮断フィルタ（波長750nm以下遮断、780nm以上透過）を設けた。また、カメラには、外部照明の影響を少なくするため同様の可視遮断フィルタを付けた。

眼は、カメラの光軸上で約1mの位置に置き、眼を左右方向に各々10、20度回転したときの瞳孔像を撮影した。撮影画像の輝度分布は、階調8ビットの画像処理装置を用いて測定した。また、照度は眼と同じ位置に光パワーメータを置き空間伝播光を測定した。照明を強くすると、瞳孔から出る光のみならず、虹彩や顔の皮膚からの反射光も強くなるが、特徴点の抽出では、背景に対する特徴点部分の強度が重要なため、ここでは、虹彩または比較的明るい顔皮膚部分に対する瞳孔の明るさ（階調の差分値）を求めた。

実験結果を図6に示す。照度とともに階調差分値は大きくなり、同図に(Dark)で示す外部照明がない場合では、 $10\mu\text{W}/\text{cm}^2$ 以下の照明でも2値化が可能であった。なお、図6で差分値のばらつきは、眼の回転に対する眼底部での反射率の変化[13]、および、撮影毎の瞳孔径の変化に起因していると思われる。

ところで、通常のインタフェース環境では、外部照明が全くない状況は考え難い。そこで、近赤外に波長域を持つ外部照明がある場合について、2値化に必要な照度を調べた。外部照明は、瞳孔以外の部分の輝度を高めるように作用するため雑音光源と言える。実験では、図5に示すように、白熱灯を使用し、これによる眼の近傍の光強度を約 $15\mu\text{W}/\text{cm}^2$ とした。この強度は可視照度計で測定すると、約15 Luxに相当し、外部照明としてはかなり暗く感ずる条件である。結果は、図6に示すように、(Dark)の場合の階調差分値より下がり、2値化には数 $10\mu\text{W}/\text{cm}^2$ の照明が必要であった。このように、2値化に必要な照明強度は外部照明の強度とともに大きくなる。

眼に照射できる赤外光の許容値については、基準が見当たらないが、紫外光に対しては、実効放射照度と許容被ばく時間との関係について、ACGIH（米国労働衛生専門官会議）の勧告が知られている[4]。それによると、上記の実験でDarkの場合に必要な照度である $10\mu\text{W}/\text{cm}^2$ に相当する紫外線エネルギーの、1日の許容被ばく時間は5分であり、照度が低くなるに従って、許容時間は指数関数的に増大する。この勧告値は、赤外光に直接適用はできないが、可視以外の光は有害と言う観点に立つと[4]、近赤外光の照度もできるだけ低くする必要がある。

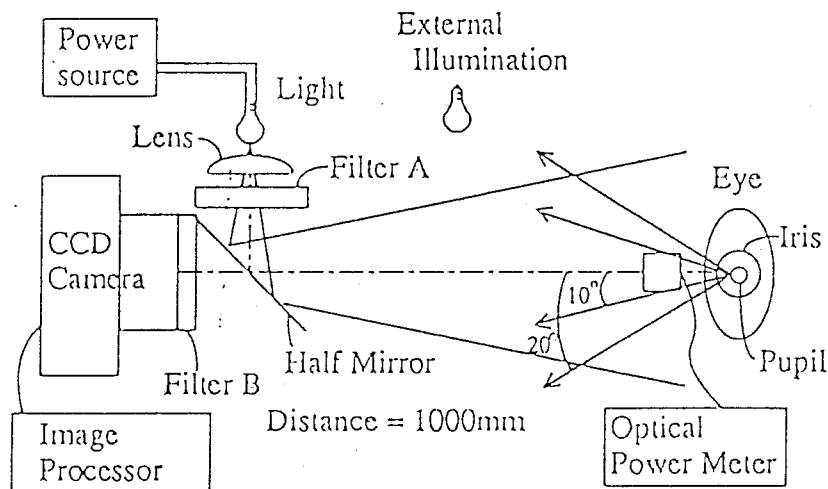


図5 瞳孔輝度の計測系

Fig. 5 Pupil brightness measurement system

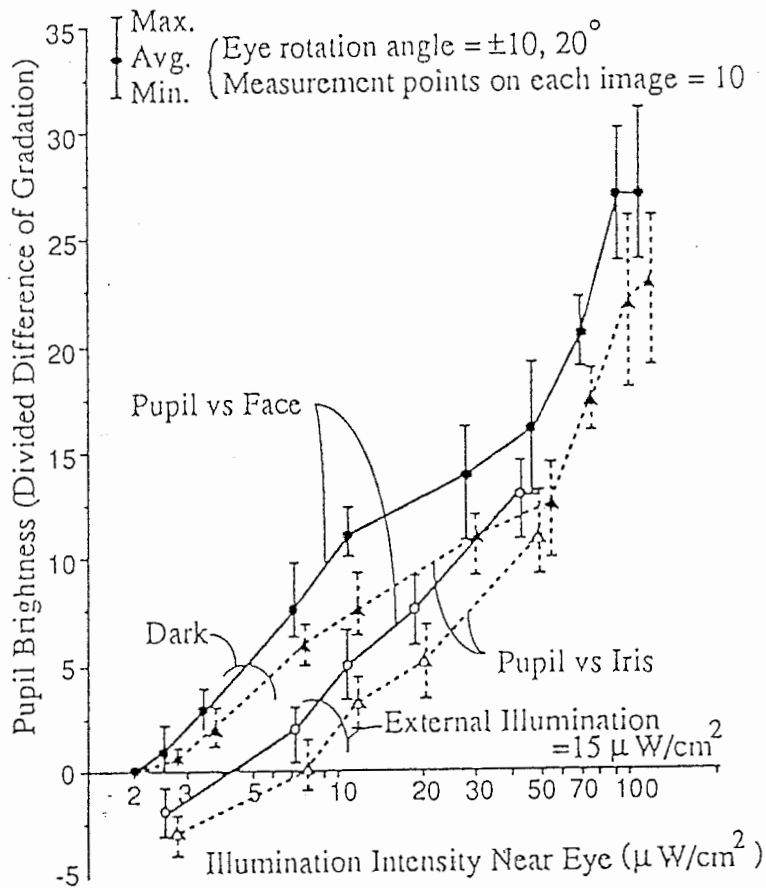


図 6 照明強度に対する瞳孔輝度の変化

Fig. 6 Pupil brightness corresponding to illumination intensity

### (b) 照明の波長と瞳孔像

(a) では、波長780nm 以上の光を全て使用した。しかし、眼底の構成要素である網膜には、光の透過率に波長特性があり、800nm から1000nmにかけて低下することが知られているため[4]、仮に、眼底からの反射率にも同様の特性があるとすると、波長を選択して照明した方が、弱い照明で効率の良い撮影が期待できる。そこで、図5において、可視遮断フィルタの代わりに、各々20~30nmの半値幅を持つ、6種類のバンドパスフィルタを用い、眼近傍での照度を一定( $10\mu\text{W}/\text{cm}^2$ )とした条件で、波長に対する瞳孔の輝度変化を測定した。

図7に結果を示す。800nm から1000nmの波長域において、波長の長さとともに階調差分値は大きく低下する。この理由は、眼底部での光の反射率、および、CCDの感度(図14参照)に、波長依存性があるためである。また、950nm 以上の波長は瞳孔の輝度にあまり寄与していないことも分かる。従って、少ない光エネルギーで効果的に撮影するためには、可視光及び波長の長い近赤外光を遮断するバンドパスフィルタを用いるのが望ましい。

### (c) 強調化処理と効果

(a) の実験で示したように、眼底反射光を捉えて瞳孔を明るく撮影する手法は、外部照明の影響がない場合には有効であるが、外部照明などによって瞳孔以外の部分が明るく撮影される場合には、その効果は低下し、単純な2値化処理が適用できない場合がある。瞳孔とその周辺との輝度差を大きくし、弱い照明でも2値化処理を可能にする一手法として、瞳孔が明るい画像と、暗い画像を同時に撮影して、2つの画像を差分処理する手法が考えられる。瞳孔を明るく撮影する照明の位置は先に述べた。また、暗く撮影するためには、カメラレンズの有効径から2~3cm以上離れたところから照明すればよい[11]。

そこで、差分処理の有効性を図8の実験系を用いて測定した。第1の照明は瞳孔を明るく撮影するためのもので、ハーフミラーを用いて光軸付近から照射した。第2の照明は瞳孔を暗く撮影するためのもので、レンズの外側にLEDを円形に設けた。第1、2の照明装置の波長は僅かに変化させておき、これらを同時に照射したときの、反射光をカメラの中で波長分離し2つのCCDに結像させることにより、同時に2種類の画像を撮影した。なお、波長が異なると、被写体の各部分の反射率が変化するため、瞳孔以外にも輝度差を生ずる部分が出る可能性があるが、撮影範囲は主に顔であり、波長の差が小さければ、瞳孔周辺部分の変化は瞳孔部分の変化に比べて十分に小さいと考えられる。従って、差分処理で瞳孔が強調できると考えた。このようなカメラ構成については、4章で述べる。ここでは、原理を確認するため、顔と眼を固定し、2種類の波長の照明を時分割で交互に駆動し、これに連動して単板CCDカメラで撮影した。

実験では、図8において、照明(I)を800nm、照明(II)を850nmとし、前記(a)の実験結果と比較するため、2つの光強度の和をパラメータとして変化させた。また、外部照明も(a)の場合と同様とした。実験結果を図9に示す。眼近傍に照射する光強度が同じ条件で(a)と比較すると、階調差分値は増加し

改善が見られた。また、 $10 \mu\text{W}/\text{cm}^2$  以下の比較的弱い照明でも抽出できることが分かる。

以上、効率的な瞳孔撮影の条件について述べたが、光強度を更に下げるとともに、撮影するときのみ照明を行うなどの手法を合わせて検討する必要がある。

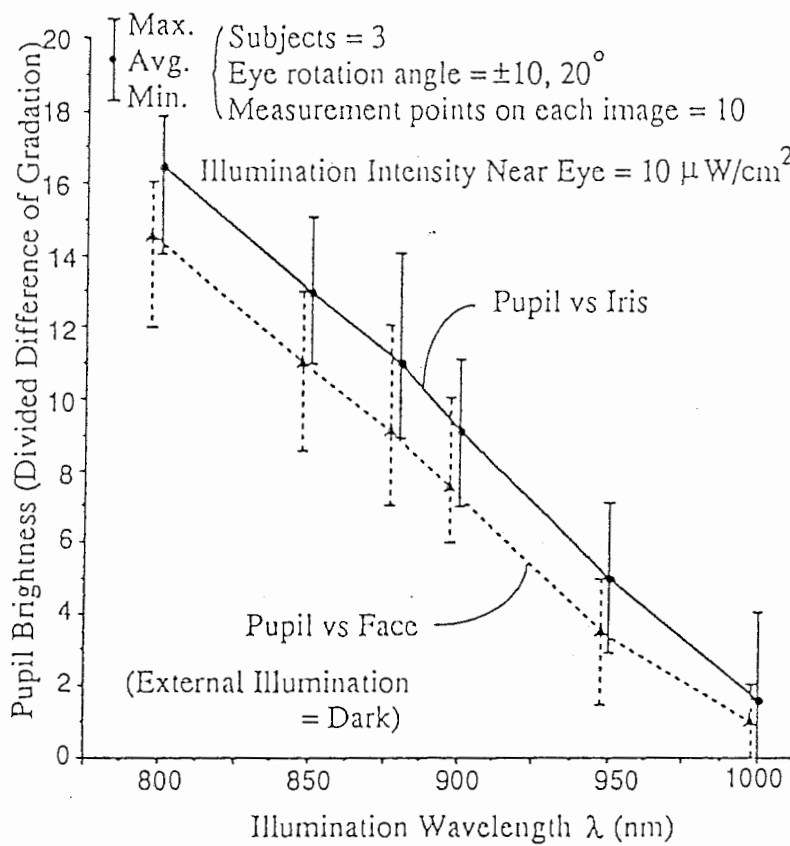


図7 照明光の波長に対する瞳孔輝度変化

Fig. 7 Pupil brightness corresponding to illumination wavelength

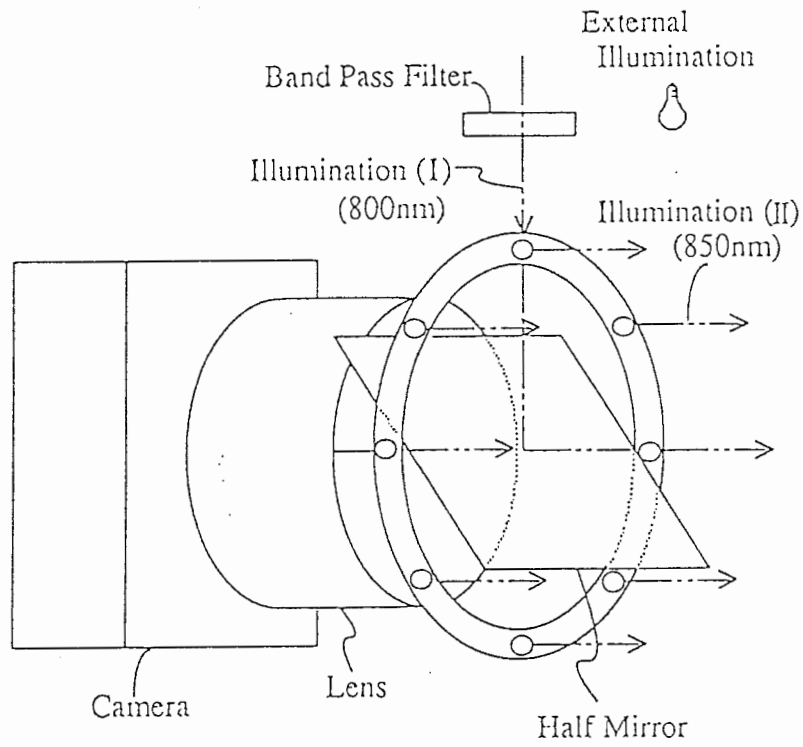


図 8 差分処理による瞳孔抽出実験

Fig.8 Pupil extraction experiment system using subtractive processing

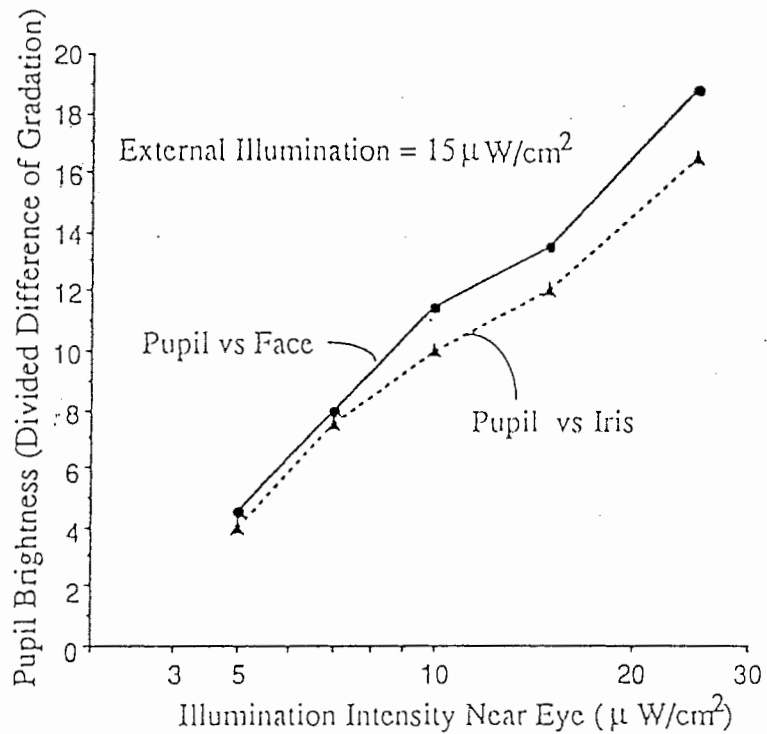


図 9 差分処理による瞳孔部の強調

Fig. 9 Pupil brightness accentuated by subtractive processing



### 3. 2 顔マーク

2. 1で述べた、顔と瞳孔を特徴点とする視線検出法では、顔を三角形でモデル化するため、顔の中で動きの少ない3つの点を抽出する必要がある。将来的には実際の顔画像から自動的に抽出することも考えられるが、現状では様々な制約があるため、便宜上、顔にマークを付け、それを特徴点とする方法を考えている。マークは瞳孔と一緒に抽出する必要があり、そのためには瞳孔と同じ原理で撮影できるマークを用いるのが効果的である。そこで、小さなガラス球の表面を一部分円形の開口部を残して塗料をぬり、開口部が眼球の瞳孔のように作用するマークを提案した[12]。このマークを眼鏡フレームに取り付けたものを、顔に装着することにより、顔の3点と瞳孔を抽出する。抽出結果については、4. 4で述べる。

### 3. 3 角膜反射像

角膜反射像は、2. 1(2)で述べたように、入射光が角膜で反射することによって、眼の中に生ずる虚像である。従って、特徴点生成のための光源が小さく、光が強ければ、これによる角膜反射像は明るい輝点として撮影されるため、2値化処理により比較的簡単に抽出できる。しかし、実際のインタフェース環境では、外部照明があり、これらも眼の中に角膜反射像を作るため、特徴点抽出の障害となることがある。このような状況は、自然光が入射している室内や、白熱灯を使用している場合のように、外部照明の波長が特徴点生成光源の波長と重なっている場合に問題となる。

そこで、特徴点生成光源により作られる角膜反射像だけを背景から分離する一手法として、偏光化した光源を用いる撮影方法を提案する。撮影原理について述べる。光源の前に偏光板を設け眼の近傍に照射すると、角膜表面は滑らかなため、偏光は崩れずに反射する。一方、角膜以外の部分、例えば、顔の皮膚などでは微細な凹凸があるため、照射された光は拡散反射し偏光は崩れる。

この性質を利用すると、反射光をカメラで捉える際、偏光成分を遮断するか、通過させるかにより、角膜反射像を含む画像と含まない画像が別々に得られたため、2つの画像の差分をとれば、特徴点生成光源による角膜反射像の輝度は強調されると考えられる。なお、偏光成分の遮断または通過の選択は、撮影光学系の光路に偏光板を設け、この偏光角を光源の前の偏光板に対して90度回転させるか、同じ角度にするかによって制御できる。

一方、外部照明は、通常、偏光性が弱いので、これによる角膜反射像の輝度は、上記の撮影系の光路に設けた偏光板の偏光角によって、その影響を受けにくい。従って、差分処理によりその輝度は大きく低下すると思われる。

以上の原理を確認するため、図10に示す撮影光学系を用いて、角膜反射像を強調化する実験を行った。原理からすると、本来は、同一光源を用いた際の反射光をカメラの中でハーフミラーなどで2つに分岐し、偏光角が各々90度異なる偏光板を介して結像させ、2つの画像の差分をとるべきであるが、ここでは、実験の都合上、偏光角が直交する2つの光源をハーフミラーを介してそれらの光軸が一致するように配置し、時分割で駆動し、これに同期させて撮影したフレーム間画像の差分をとった。

光源には、850nmの波長を持つLEDを用い、LED(1)、(2)の前には、偏光角が直交する偏光板P1、P2を設けた。また、カメラの前には、P1と同じ偏光角の偏光板P3を設けた。図11は、使用したLEDの波長特性、偏光板の透過特性（2枚の偏光板の偏光角が直交している場合が曲線C、偏光角が同じ場合が曲線B）、及びCCDカメラの感度特性である。

実験結果を図12に示す。(i)(ii)は、各々、LED(1)、(2)を照射したときの撮影画像である。前者には角膜反射像が含まれるが、後者には含まれず、ほぼ原理通りの結果になっている。(iii)は、2つの画像の輝度を差分処理し、ヒストグラムを求めたものである。角膜反射像以外の部分は輝度が低下し、相対的に角膜反射像が強調されている様子が分かる。

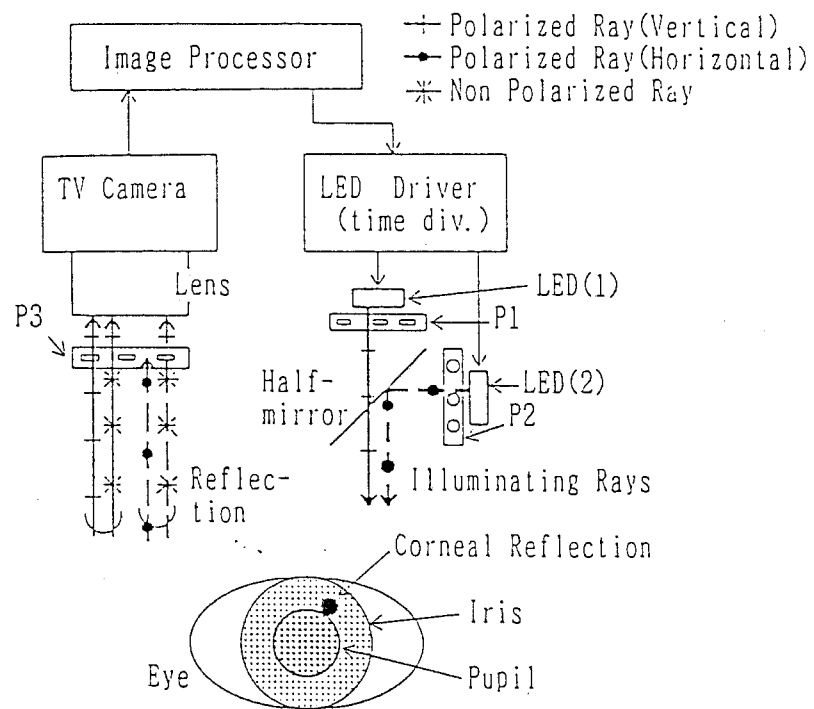


図10 角膜反射像の抽出実験

Fig.10 Corneal reflection extraction experiment.

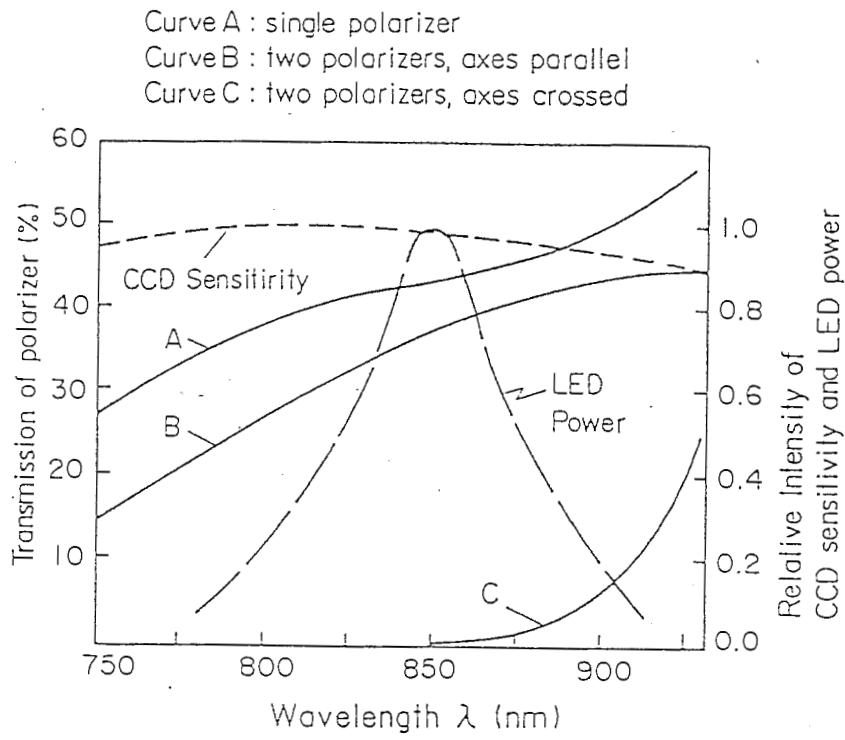
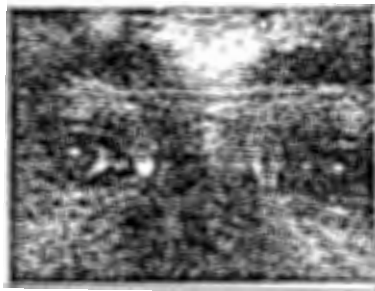
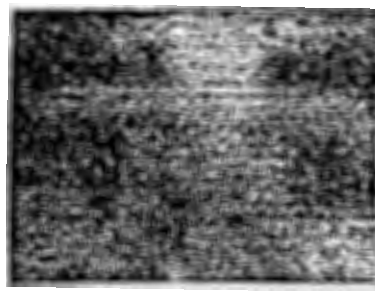


図 11 照明スペクトルに対する偏光板の特性

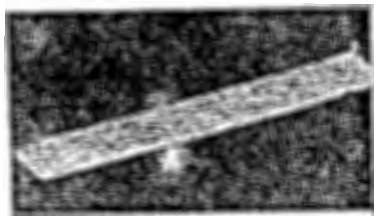
Fig.11 Polarizer characteristics corresponding to illumination spectram



(i) Image using LED(1)  
 (Corneal reflections  
 Appear)



(ii) Image using LED(2)  
 (Corneal Reflections  
 Disappear)



(iii) Histogram of  
 Subtraction Image

図 12 角膜反射像の抽出結果

Fig. 12 Results of corneal reflection extraction

#### 4. 近赤外 3 板式 CCD カメラの試作

3. で述べた方法に基づき、瞳孔、顔マーク、及び、角膜反射像の各々を協調して撮影するため、図 13 に示すような近赤外 3 板式 CCD カメラを試作した。

##### 4. 1 照明系

本撮影法では、何れの特徴点の場合でも、特徴点部分の輝度のみが異なる 2 種類の画像を撮影し、差分処理して強調する。

瞳孔と顔マークの撮影では、3. 1 (C) で述べたように、2 か所からの照明による画像を同時に、且つ分離して撮影するため、波長の僅かに異なる近赤外照明を用いた。図 13 において、照明(1)は、波長 850 nm で、光軸近くに設け瞳孔を明るく撮影するために用いた。また、照明(2)は、波長 950 nm で、瞳孔を暗く撮影するためレンズ有効径から 2 cm 余り離して設けた。照明装置の波長分布特性を図 14 に示す。また、角膜反射像の撮影では、照明(1)を光源として使用するため、この前に偏光板を設けた。

##### 4. 2 プリズム光学系

2 種類の照明による反射光を 3 つに分離し、結像させる光学系を図 13 に合わせて示す。反射光は、レンズを通過しプリズム光学系に入る。波長  $\lambda_2$  ( $=950\text{nm}$ ) の光は、第 1、第 2 のプリズム境界面で分離され、CCD 3 の撮影面に結像する。波長  $\lambda_1$  ( $=850\text{nm}$ ) の光は、第 2、3 のプリズムの間に設けたハーフミラー膜によって分離され、一方は直接 CCD 2 に、他方は偏光板を介して CCD 1 に結像する。ここで、偏光板は、照明(1)の偏光と直交するように設定した。また、偏光板 1 枚の透過率は、図 11 に示したように 40~50% である。偏光板を通過すると、光は弱くなるため、ハーフミラー膜の分割比を同図のように 1:2 とし、CCD 1 と CCD 2 の撮像面での光強度が同程度になるようにした。プリズム光学系の各通路 (CH1, CH2, CH3) の光透過特性を図 14 に合わせて示す。CH1 と CH2 は同程度、CH3 は約 3 倍になっている。また、同図には、レンズの透過特性、CCD の相対感度特性を合わせて示す。これらは、照明の波長域でほぼ平坦な特性となっている。

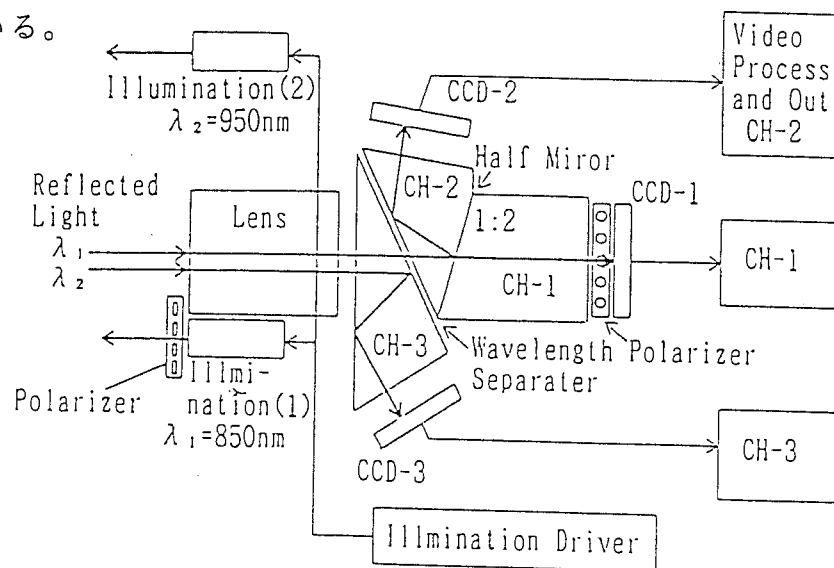


図 13 3 板 CCD 式近赤外カメラの構成

Fig. 13 Infrared 3-CCD camera system

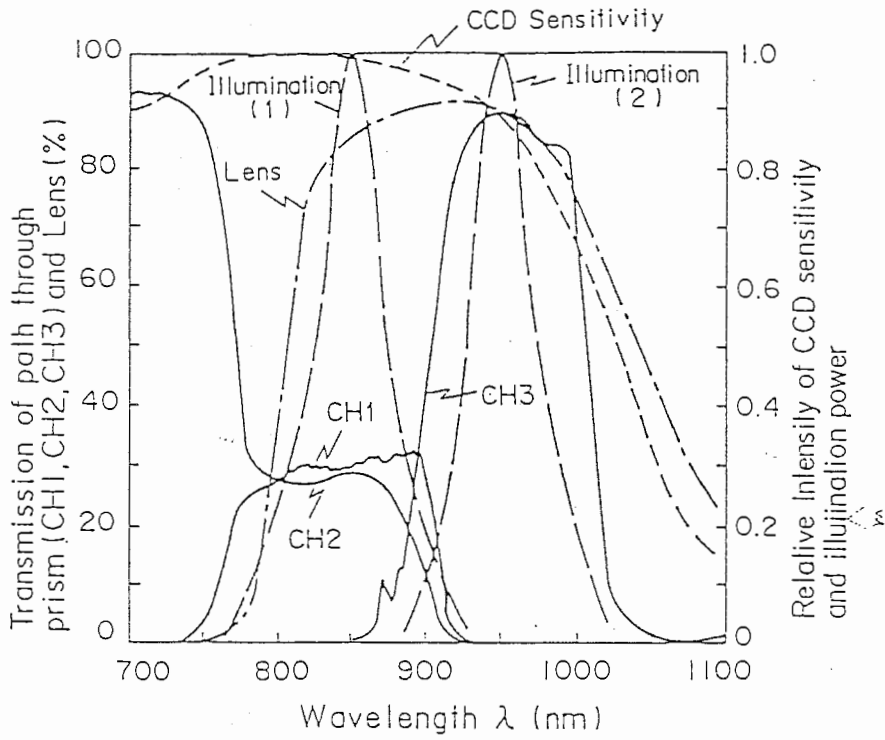


図14 カメラ光学系の特性

Fig.14 Characteristics of camera optical system

表2 近赤外CCDカメラの特性  
Table 2 Characteristics of Infrared 3-CCD Camera System

Items	Characteristics
Input Wavelength	$\lambda_1=850\text{nm}$ , $\lambda_2=950\text{nm}$
Separation Wavelength	$\lambda_s =900\text{nm}$
Image Pick-up Device	Frame Transfer Type 2/3" CCD
Picture Elements	800(H) $\times$ 490(V)
Polarizer	Film Type
Resolution	Horizontal; 560 TV Lines
TV Field Scanning	2:1 Line Interlace
Electronic Shutter	1/100, 1/250, 1/500, 1/1000(s)

#### 4. 3 駆動系とカメラ特性

図13において、各CCDの出力信号は、NTSCビデオ信号に変換される。ここで、プリズムのCH1, CH2の透過率は前述のように低いため、これらのチャンネルの信号は増幅し、CH3と同レベルになるようにした。試作したカメラの主な仕様を表2に示す。

このカメラを用いて、撮影実験を行った結果、照明装置(1)によるCH1, CH2の撮影画像では、原理通り瞳孔は明るくなった。また、両者を比較すると、CH1の画像では、偏光板の効果により、角膜反射像の強度が低下した。2つの画像の差分処理により角膜反射像は強調された。一方、照明装置(2)によるCH3の画像では、瞳孔は暗く撮影された。2つの照明装置の波長は約100nm異なるため、眼の周辺の2つの画像の輝度分布はやや異なるが、瞳孔部分の輝度差が最も大きいため、CH2, CH3の画像の差分処理により瞳孔は強調された。

#### 4. 4 実時間画像処理

以上のように、差分処理により、比較的簡単に特徴点が強調化できるため、高速な2値化処理と中心計算処理を組み合わせれば、特徴点位置の実時間検出ができると考え、図15に示すパイプライン画像処理装置を用いて実験を行った。

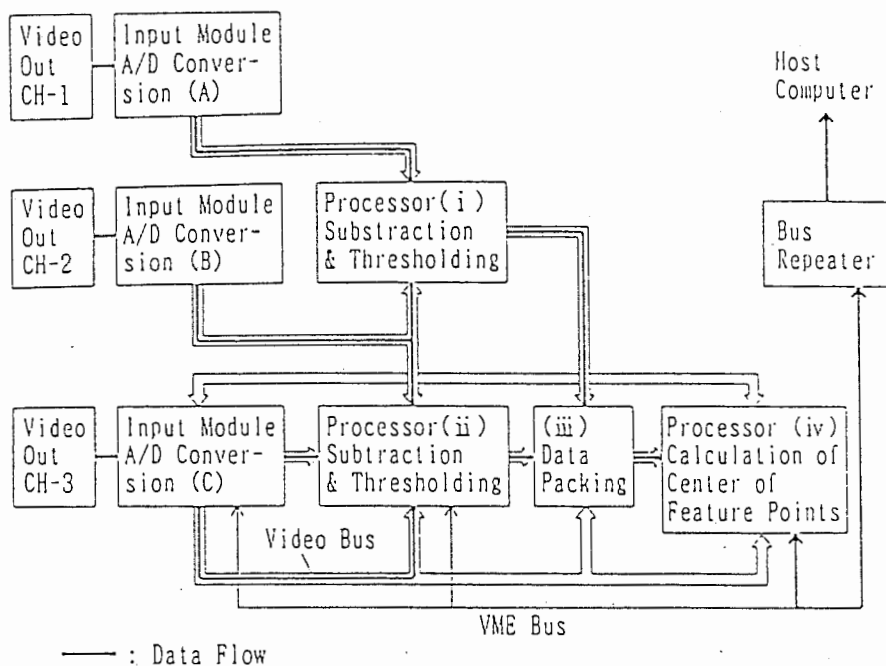


図15 パイプライン画像処理装置の構成

Fig. 15 Block diagram of real time image processing apparatus

装置の概要について述べる。入力モジュール(C)は、パイプライン処理のための基本モジュールで、タイミング信号を作り、各モジュールに配信する。各モジュールはこの信号を媒体として画素単位で並列動作する。このため、画像内、画像間に係わらず、簡単な演算であれば、原理的には、～数10画素の計算遅延で処理が進行する。

データの流れを図中太線に示す。前述の3板式CCDカメラの各ビデオ信号は、入力モジュール(A),(B),(C)に取り込まれる。ここで、これらのモジュールは、カメラの出力信号に同期して、ビデオ信号をA/D変換し、この画素データをVideo Busを経由して順次各演算処理モジュールに送る。Processor(i)は、CH1とCH2の画像を、Processor(ii)は、CH2とCH3の画像を、各々差分処理および2値化処理する。これにより、各特徴点は画面中で各々固まりとして抽出される。Processor(iii)は、(i)、(ii)によって作られた2つの2値化画像をパッキングし、Processor(iv)は、特徴点を構成する画面中の複数の画素位置を平均し、特徴点の中心点として出力する。

このように、本装置では、画像の取込みに同期して、A/D変換が済んだ画素データは、複数のモジュールを通過する間に処理されるため、画像処理による遅延時間は、1枚の画像の取込みに必要な時間33msにパイプライン処理の遅れ時間を加えたものになる。

図16は、4つの特徴点(顔マーク3点、瞳孔1点)に対して、差分処理、2値化処理、および中心位置計算処理を行った例である。遅延時間を測定したところ、パイプライン処理の遅れは約5ms程度、画像の取込みと合わせても38ms程度であった。この特徴点位置データは、ホストコンピュータで視点、視線に変換されるが、この時間を加えても2.2(2)で述べた60msの目標は満たされるものと思われる。

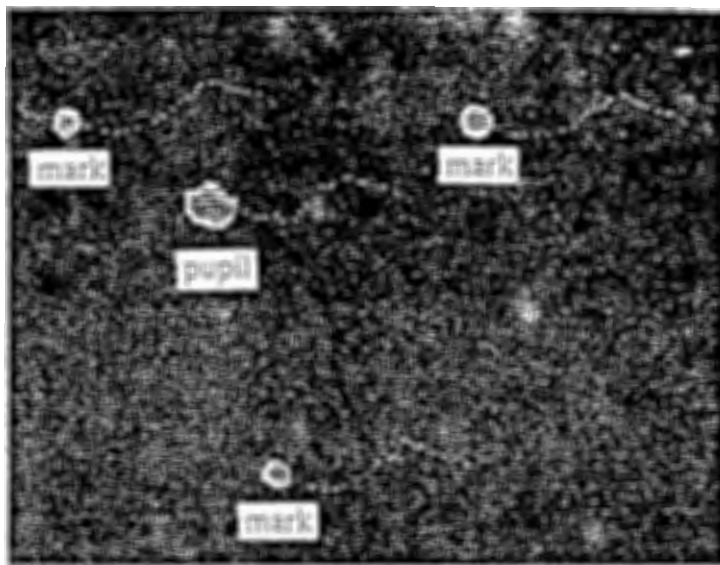


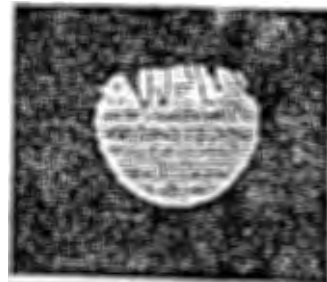
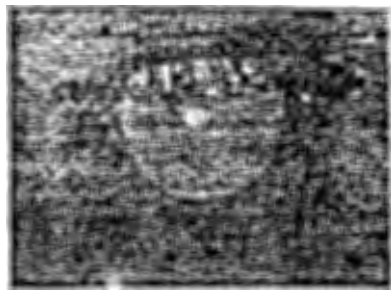
図16 顔マークと瞳孔の実時間抽出結果

Fig. 16 Result of extracting 3 marks on the human face and pupil

## 5. 中心位置計算の高精度化

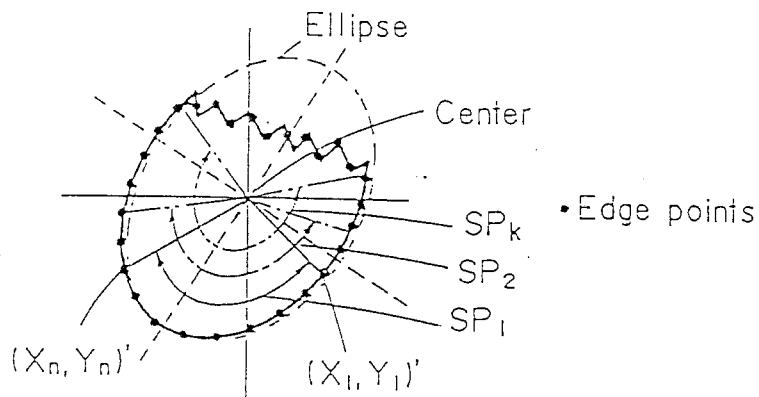
### 5.1 楕円近似による輪郭データの補正方法

視点・視線の検出では、特徴点の高速な抽出とともに、撮影画面中でその中心位置を正確に求めることが重要である。4.4で述べた実時間処理では、特徴点を構成する画素位置を平均して、特徴点の中心位置とした。特徴点が完全な形で抽出できる場合には、この方法でも誤差は少ないが、特徴点の中でも瞳孔は、図17のように、瞼や睫毛により遮られ、その形状が部分的に欠落する場合が多いため、単純な平均では誤差が大きくなり、2.2で述べた目標は困難である。そこで、中心位置計算の精度を上げるため、楕円近似による輪郭データの補正を試みた。



(a) Picked up pupil image

(b) Result of thresholding process



(c) Ellipse fitting method

17  
 図 17 瞼などによる瞳孔の部分的欠落と楕円近似による補正方法

17  
 Fig. 17 Partial lack of pupil caused by eye lid and correcting method using ellipse fitting.



本検出方法では、瞳孔に限らず、特徴点は円を想定しているため、カメラで撮影した画像には、楕円を当てはめることができる。抽出された特徴点の輪郭データを周に沿って分割し、楕円で近似できるデータのみを用い中心を推定する。以下、最小自乗法による楕円近似の方法を説明する。

図17(c)において、●印は輪郭の点データである。この中の部分的なサンプル点データを、 $SP_1: \{ (x_1, y_1)' (x_2, y_2)' \cdots (x_n, y_n)' \}$  とすると、これを楕円で近似する場合の目的評価関数  $S$  は次式で与えられる。

$$S = \sum_{i=1}^n \{ x_i^2 + a x_i y_i + b y_i^2 + c x_i + d y_i + e \}^2 \quad (1)$$

$S$  を最小とするための条件は、右辺を  $a, b, c, d, e$  の各々で偏微分した値が0であり、これを整理すると、

(2) 式になる。

$$\begin{aligned} a [x^2 y^2] + b [x y^3] + c [x^2 y] + d [x y^2] + e [x y] &= - [x^3 y] \\ a [x y^3] + b [y^4] + c [x y^2] + d [y^3] + e [y^2] &= - [x^2 y^2] \\ a [x^2 y] + b [x y^2] + c [x^2] + d [x y] + e [x] &= - [x^3] \\ a [x y^2] + b [y^3] + c [x y] + d [y^2] + e [y] &= - [x^2 y] \\ a [x y] + b [y^2] + c [x] + d [y] + e [1] &= - [x^2] \end{aligned} \quad (2)$$

ただし、 $[x^2 y^2] = \sum_{i=1}^n x_i^2 y_i^2$ 、他も同様である。

(2) 式より求めたパラメータを  $A, B, C, D, E$  とすると、求める楕円関数は、(3) 式となる。

$$x^2 + Axy + By^2 + Cx + Dy + E = 0 \quad (3)$$

ここで、中心座標の算出のため、(4) 式で示す座標変換を(3) 式に行う。

$$x = X + x_0, \quad y = Y + y_0 \quad (4)$$

これが、楕円であるためには、 $X, Y$  の一次項=0 の条件が成立する。この条件式を  $x_0, y_0$  について解くと、

中心座標の推計値は、(5) 式で与えられる。

$$x_0 = - \frac{Ay_0 + C}{2}, \quad y_0 = \frac{AC - 2D}{4B - A^2} \quad (5)$$

ここで、サンプル点の採り方について簡単に説明する。推計の精度を上げるためには、瞼などの影響を受けないできるだけ多くのサンプル点を用いるのが望ましい。図17(c)の例では、 $SP_k$  の範囲のサンプル点となる。輪郭データの中からこれらの点だけを選択する方法の一例を以下に示す。

① 瞳孔は、図17(a)に示すように、上の方が瞼の影響を受けやすく、横または下の方は比較的影響を受け難いと言う性質がある。そこで、影響を受け難い部分に初期のサンプル点を設け(図18(c)の $SP_1$ )、上記方法で楕円を当てはめる。

② この楕円と輪郭の各点との距離を順次求め、所定の距離内で、前のサンプル点に連続する点を追加し、新しいサンプル点とし、再度楕円を求める( $SP_2$ )。

③ ②の処理を追加する点がなくなるまで繰り返し、このときの楕円中心を推計

結果とする。

以下では、輪郭データの中から瞼などの部分を除去できたとして、楕円近似による中心位置検出精度を評価する。

## 5. 2 シミュレーション実験

楕円近似を適用して中心位置を求める場合、輪郭の各点データに含まれるピクセル誤差が中心位置推計の精度に影響する。そこで、先ず、ピクセル誤差をモデル化する瞳孔像を作成した。

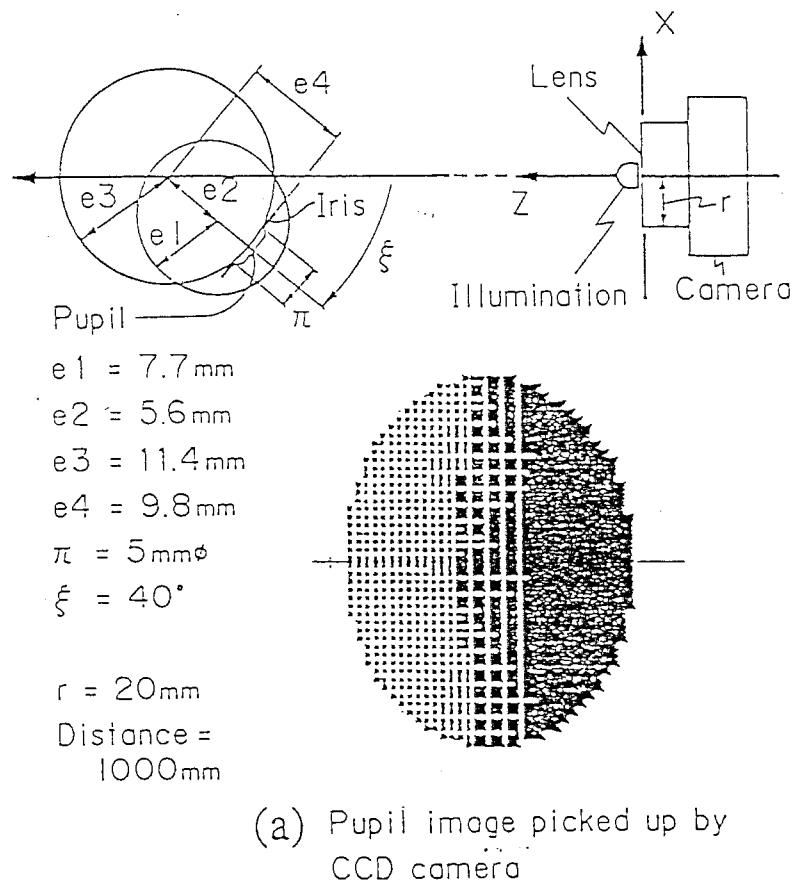


図 18 瞳孔像の計算条件と計算結果

Fig.18 Calculating condition and result of  
 picking up pupil image

### (1) 瞳孔像の作成

図 18 に眼球と撮影系のモデルを示す。カメラの光軸近傍に光源を設け、この照明による眼底での反射光をカメラで捉えたときの瞳孔像はレイトラッキングによって求めることができる[12]。この瞳孔像を CCD の画素で分割し、強度分布を計算すると、図 18 (a) になる。なお、同例では瞳孔径が CCD 上で 40 画素程度に撮影される場合を示している。この大きさは、撮像素子の画素数を  $500 \times 500$  程度とすると、画面の  $1/10$  程度に相当する。瞳孔径に対応する画素数が多い程、中心位置の検出精度は高いが、画面上での特徴点の動きを考えると、画面の  $1/10$  程度の大きさは拡大して撮影できる限界に近いと思われる。

(2) 瞳孔像の欠落と精度の関係

図18(a)を2値化して瞳孔部分を抽出し、瞳孔が上部から下部に向かって欠落する場合について、中心位置を求めた。結果の一例を図19に示す。●及び○が2値化処理で抽出される画面上の瞳孔部分の座標点である。(i)、(ii)は、各々瞳孔の20%、80%が欠落した場合である。●の輪郭点座標を用いて楕円近似し、中心点位置を推計した結果を▲、また、瞳孔部分の各点座標を単純に平均して求めた結果を△で示す。

図20は、瞳孔の欠落面積をパラメータとして、モデルで与えた真の中心点(×)と、推計、または、平均により求めた中心点との距離を誤差として求めた結果である。平均処理では、瞳孔面積が小さくなるに従って誤差が大きくなり、表1に示した目標の達成は困難なことが分かる。一方、楕円近似による推計を適用すると、瞳孔面積が30%以上の場合については、0.2mm以下の誤差となっている。この誤差が、そのまま空間位置計測誤差に対応する訳ではないが、ステレオ画像計測におけるカメラの較正を精度良く行うことによって、近い値になるものと思われる。従って、表1の目標に対して考察すると、特徴点を拡大して撮影し、欠落を楕円近似で補正することができれば、所期の用途に適用可能な視点・視線検出が実現できると思われる。また、本シミュレーションでは、瞳孔輪郭の抽出に簡単な2値化を用いているが、アナログ情報を利用する方法も知られているため[14]、併用により更に高精度化が期待できる。

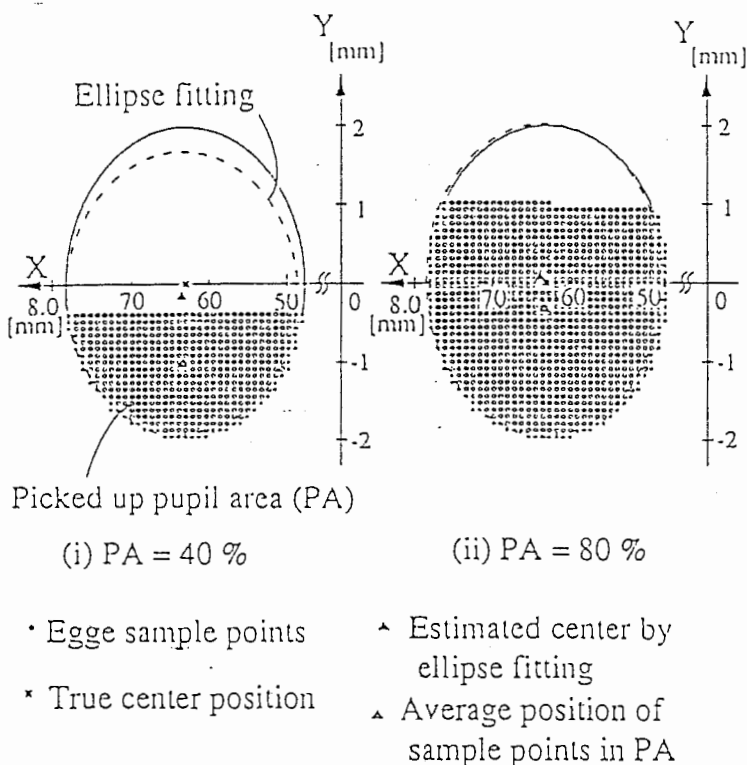


図19 瞳孔の中心位置計算例

Fig. 19 Calculation of pupil center position

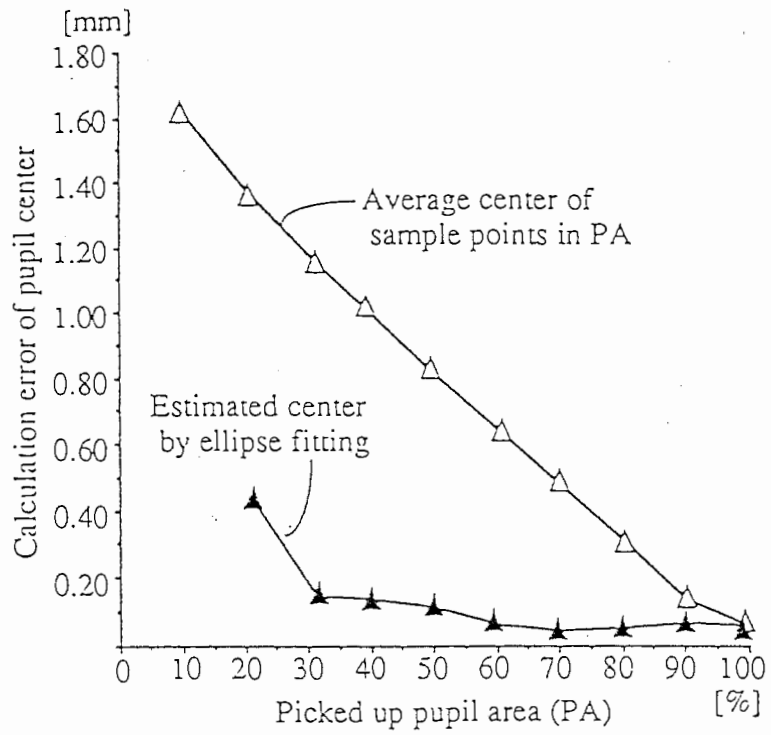


図 20 瞳孔撮影面積に対する瞳孔中心位置計算誤差

Fig. 20 Relationship between picked up pupil area and pupil center calculation error

表 1 視点・視線検出の用途と目標特性

用途	特徴点の検出精度		検出速度
	相対位置	絶対位置	
運動視CGのための視点検出	< 1.0 mm	—	< 60 ms
指示入力のための ①顔と瞳孔を用いた視線検出	< 0.4mm	< 5.0 mm	< 200ms
②瞳孔と角膜像を用いた視線検出	< 0.15mm		

## 6. むすび

画像処理による視点・視線の検出において、特徴点抽出に要求される目標特性を設定し、これを実現するための高速、高精度な画像処理方法について検討した。その結果、以下のことを明らかにした。

- ① 眼底反射光を捉える瞳孔抽出法では、瞳孔像の明るさは、照明の波長に依存する。近赤外では波長が長くなるに従って反射光が弱くなり、瞳孔像の輝度には寄与しないため、950nm以上の光は遮断するのが望ましい。
- ② 瞳孔抽出に必要な照度は、外部照明が強くなるに従って高くなる。近赤外光の眼への影響を考慮しすると、常に弱い照明で抽出できることが望ましく、その一手法として差分処理が有効であった。
- ③ 角膜反射像の生成に偏光照明を利用し、偏光成分を遮断した画像と、通過させた画像を同時に撮影し、差分処理すると、角膜反射像を背景から効率よく抽出できる。
- ④ パイプライン画像処理装置を用いることにより、差分処理、2値化処理による特徴点の抽出、および、その中心点位置の計算を40msec以下で行うことができる。
- ⑤ 特徴点の中心位置検出精度を上げるため、楕円近似を適用し、シミュレーション実験を行い、視線検出に必要な精度が得られる見通しを得た。

## 参考文献

- [1] 伴野 明, 岸野文郎: " 臨場感通信会議におけるヒューマンインタフェース技術 ", 人工知能学会誌, 6, 3, pp. 358-369(1991.05).
- [2] 山口博幸, 伴野 明, 岸野文郎: " 視線検出を利用する広視野高精細表示方法の検討 ", 信学論, C-II, J73-C-II, 11, pp. 721-732(1990.11).
- [3] 伴野 明, 飯田宗夫, 小林幸雄: " 非接触視線検出のための特徴点出法 ", 信学技報, PRU88-73 (1988.11).
- [4] 関 亮: " 紫外・赤外線の目に対する作用効果と安全 ", 照明学会雑誌, 61, 11, pp. 636-641(1977).
- [5] 伴野 明: " ステレオ画像計測を用いた視線検出法の提案 ", テレビ技報, Vol. 14, No. 36, pp. 19-24 (1990.06).
- [6] 大村和典, 伴野 明, 小林幸雄: " 単眼視画像による顔の向き検出法の指示入力への応用 ", 信学論, J72-D-II, 9, pp. 1441-1447 (1989.09).
- [7] 伴野 明, 佐藤隆宣, 岸野文郎: " 透視投影画像生成表示における視点位置の微小変化と画質特性 ", 1991年信学秋季全大, (1991.09).
- [8] Young. L and Sheena. D: " Survey of eye movement recording methods ", Behav. Res. Methods & Instrum., vol. 7(5), pp. 397-429(1975).
- [9] 星野春男, 比留間伸行, 山田光穂, 福田忠彦: " 運動視差利得を変化させたときの視覚効果 ", 信学論 (D-II), J73-D-II, 2, pp. 224-231(1990.02).
- [10] 山田光穂, 福田忠彦: " 画像における注視点の定義と画像分析への応用 ", 信学論, J69-D, 9, pp. 1335-1342 (1986).
- [11] Hutchinson. T. E, White. K. P, Martin. W. N, Reichert. K. C and Frey. L. A: " Human-computer interaction using eye-gaze input ", IEEE Trans. Systems, Man & Cybernetics, 19, 6, pp. 1527-1534(1989-11).
- [12] 伴野 明: " 視線検出のための瞳孔撮影光学系の設計法 ", 信学論 (D-II), J74-D-II, 6, pp. 736-747(1991.06).
- [13] Hatada. T and Ohzu. H: " Optical Modulation Characteristics and Chromatic Fundus Reflectometry of the Living Human Eye ", Memoirs of the School of Science & Engineering Waseda Univ., No. 35, pp. 103-117(1971).
- [14] 笠井 健, 東野直巳, 辻 三郎: " CCDイメージセンサを用いた眼球運動計測法とその測定分解能 ", 信学論(D), J65-D, 2, pp. 274-281 (昭57-02).