

複数カメラを利用した広域遠隔視線計測

岡本 康太郎^{†,††} 内海 章[†]
萩田 紀博[†] 高橋 和彦^{††}

複数のカメラを利用し従来法に比べて広い頭部姿勢範囲について視線計測を可能にする手法を提案する。角膜反射やステレオ計測を利用した手法の問題点である観測距離の制約を緩和する手法として単眼カメラによる視線計測手法が提案されている。しかし、同手法には顔向きの変化に関しては依然として強い制約があり、広い空間を注視対象とした人の視線の動きを計測するためには多数のカメラが必要であった。提案手法では、顔面上の特徴点および左右の目領域画像の観測がそれぞれ異なるカメラで行われることを許容し、それらの観測を組み合わせることで頭部の位置・姿勢および視線方向の推定を行うことで、少数のカメラによって大きな顔向き変化に対応できる視線計測システムを構築する。実験では2台のカメラを用いた実装によって、約120度の顔向き変化への追従、平均誤差8度以内での視線推定精度が得られることを確認した。

Gaze Tracking in Wide Area Using Multiple Camera Observations

KOTARO OKAMOTO,^{†,††} AKIRA UTSUMI,[†] NORIHIRO HAGITA[†]
and KAZUHIKO TAKAHASHI^{††}

We propose a multi-camera-based gaze tracking system that provides a wide observation area. In previous studies, gaze tracking system using single camera exits the limitation regarding face orientation. In our system, Each facial feature and eye region image can be observed by different cameras, and in contrast to stereo-based systems, no shared observations are required. This feature relaxes the geometrical constraints in terms of head orientation and camera viewpoints and realizes wide availability of gaze tracking with a small number of sensors(cameras). In experiments, we confirmed that our implemented system can track head rotation of 120° with two cameras 120° apart. The gaze estimation errors are less than 8° in average.

1. はじめに

視線情報にはユーザの意図や興味対象などの情報が含まれているため、マーケティングや広告、コミュニケーション支援など様々な用途で有用であり、視線計測を利用した研究が幅広く進められている。視線計測の手法としては、角膜反射を利用したものが一般的であるが、これらの手法では精度の高い計測が行える一方で、原理的に眼球と装置を近接させる必要があり、遠距離からの計測を行うことは難しい。また、これらの手法では、ユーザが特定の参照点を注視するなど事前にキャリブレーションを行う必要があるという問題があった。そのため、視線計測の用途は実験環境下で

の特定の被験者を対象とした計測に限られており、日常生活や公共の場において不特定多数のユーザに対して視線情報を取得することは困難であった^{1),2)}。

それに対して、近年通常の画像観測を利用した視線計測の研究が盛んに行われている^{3)~5)}。なかでも眼球の3次元モデルを用いて眼球中心と虹彩(瞳孔)を結ぶ方向として視線方向を推定する手法^{6)~9)}は、特別な照明を必要とせず、従来の角膜反射による視線推定に比べて格段に低い解像度で視線計測を行うことができるという利点がある。特に、単眼カメラによる計測手法は、一定以上の解像度で観測ができればカメラとの距離に関わらず視線計測が行えることから観測距離の制約を大幅に緩和することが可能となる。また、ユーザに意識させずに自動的にキャリブレーションを行う手法も提案されており⁹⁾、これらの手法により、日常環境において視線を計測することが可能となった。

一方で、単眼カメラによる手法では、視線計測に必要な画像特徴を観測するために、カメラに対するユー

[†] ATR 知能ロボティクス研究所

ATR Intelligent Robotics & Communication Laboratories

^{††} 同志社大学大学院

Doshisha University

ザの頭部姿勢の変化が比較的狭い範囲に限定される．そのため，広域での視線計測を実現するためには，対象者を様々な方向から観測するために多くのカメラを設置する必要がある．このことは，現実の場面における視線計測の利用を妨げる要因となっていた．

本論文では，上記問題を解決するため，顔面上の特徴点，および目領域画像の観測がそれぞれ異なるカメラで行われることを許容し，それらの観測を組み合わせることで頭部の位置・姿勢推定および視線方向の推定を行う手法を提案する．これにより，少数のカメラによって大きな顔向き変化に対応できる視線計測システムを構築することが可能となる．以下では提案手法について説明し，実験により手法の有効性を明らかにする．

以下に本論文の構成を示す．次節で我々の提案手法の特徴について述べ，3節で複数カメラを利用した視線計測手法の詳細について説明する．4節では提案手法の有効性を確認するために行った実験について述べる，5節で提案手法を利用したアプリケーションの可能性について考察し，6節で本論文をまとめる．

2. 提案手法の特徴

前節で述べた通り，近年通常のカメラによる観測によって人の視線を計測する手法が提案され^{(6)~(9)}，視線計測時の観測距離の制約を大幅に緩和することが可能となった．しかしながら，一方で観測方向に関しては依然として大きな制約が残っており，現実場面における視線計測の利用を妨げている．

図1は従来法を持つ観測時の制約について説明している．図1(A)に示す角膜反射に基づく手法では，赤外照明光の到達距離の制約のために近距離（通常1m以下）での観測が必須となる．これに対して，図1(B)に示す通常の画像観測を利用した手法では，観測距離の制約が大幅に緩和される．特に単眼カメラによる視線計測では一定以上の解像度で顔および目領域が観測できれば視線計測が可能であり，カメラとの距離に関わらず視線計測を行うことができる．しかしながら，従来の手法では，視線計測に必要な画像特徴を観測するためにカメラに対するユーザの頭部姿勢の変化が比較的狭い範囲に限定されるため，頭部姿勢が大きく変化する場面で継続的に視線計測を行うためには，対象者を様々な方向から観測するために非常に多数のカメラを設置する必要がある（図1(C））．

これに対して，提案手法では図1(D)に示すように，視線計測に必要な画像特徴の観測を複数のカメラに分担させる．顔の特徴点および眼球領域がいずれかのカ

メラで観測できれば視線を計測することが可能であり，ユーザの広範囲の顔向き変化に対応することができる．図2は提案手法による視線計測の適用例であり，2台のカメラによって対象人物を異なる方向から観測している．ここにみられるように，左の画像では被験者の顔の右側の特徴点および右目領域，右の画像では被験者の左側の特徴点および左目領域がそれぞれ観測されている．提案手法では，各カメラから計測した顔の特徴点を用いて頭部の位置・姿勢を推定し3次元の眼球中心位置を求める．また，各目領域画像から虹彩位置を検出し，眼球中心と虹彩中心を結ぶ方向として視線方向を推定している（図中，黄線は推定された視線方向を示す）．これにより被験者の顔向き変化が大きい場合にも視線計測が可能となる．複数カメラを利用した物体追跡については，これまでも広く検討が行われている^{(10),(11)}．提案手法では，複数カメラによる物体追跡手法を視線計測に適用する．

次節では，提案手法における処理の詳細について説明する．

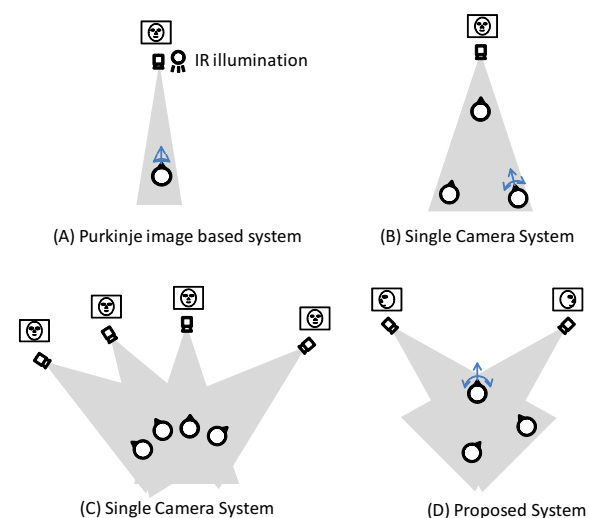


図1 手法による計測範囲の違い

Fig. 1 Difference of measurement distance and orientation



図2 提案法による視線推定の例

Fig. 2 Example of gaze estimation with the proposed method

3. 複数カメラによる視線計測処理

本節では本論文で提案する複数カメラを利用した視線計測手法の処理の詳細について述べる．図 3 にシステムの処理の流れを示す．本システムは視線方向を眼球中心と虹彩中心を結ぶ 3 次元ベクトルとして推定する．ここで虹彩はカメラによって観測可能であるが，眼球中心は直接観測することができない．そこで各カメラで得られる観測画像からユーザの顔面上の特徴点を検出し，顔形状の 3 次元モデルと顔特徴点の 2 次元観測座標から頭部の位置・姿勢を推定し，間接的に眼球中心の 3 次元座標を推定する．顔形状の 3 次元モデル（顔部品の 3 次元配置）と眼球中心位置の関係は，複数フレームの観測結果から 9) に示す手法によりユーザに意識させることなく自動的にキャリブレーションすることができる．一方で，頭部位置・姿勢の推定結果に基づいて，左右それぞれの目領域を観測可能なカメラを選択し，虹彩抽出を行う．選択された各カメラで抽出された虹彩領域について楕円形上の当てはめにより 3 次元の虹彩中心位置を決定し，眼球中心位置と虹彩中心位置から視線方向を算出する．ユーザの顔画像が 1 台のカメラのみによって観測される場合は，上記処理は単眼視線計測手法に一致する．

上記の処理を実現するためには，以下に挙げる各項目についての検討が必要となる．

- 顔の 3 次元モデル
- 顔，顔特徴点の決定
- 頭部方向の推定
- 虹彩の抽出
- 視線方向の推定

次節以降では，これらの各項目について説明する．

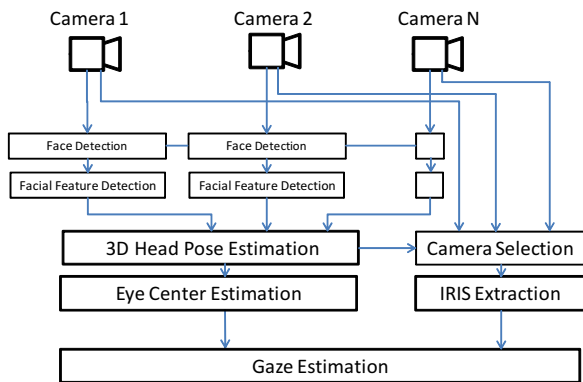


図 3 処理の流れ
Fig. 3 Process flow

3.1 顔の 3 次元モデル

顔面上の各特徴点の重心位置を原点として顔の正面方向を Z 軸の正方向とする座標系 ($X = 0$ に対して左右対称) を定義し，複数の被験者についての観測データに基づいて顔の 3 次元モデル（顔部品の配置）を生成した．

今回の実装では，顔面上の特徴点として 6 点（両目尻，目頭と口の両端点）を利用する．ここで p 番目の特徴点の 3 次元位置を (X_p, Y_p, Z_p) とする ($1 \leq p \leq 6$) ．

3.2 顔，顔特徴点の決定

顔検出および顔特徴点の検出については，それぞれ広く用いられている Haar-like 特徴量を用いた顔検出アルゴリズム¹²⁾ および，Gabor 特徴量を利用した顔特徴点抽出¹³⁾ を利用した．なお，より広い範囲の頭部角度に対応させるために，顔特徴点モデルとして横顔データを含む複数の頭部姿勢における観測データを用いた．

3.3 頭部方向の推定

上記の顔の 3 次元モデルを用いて顔の 3 次元位置，姿勢を決定する．頭部の位置，姿勢を R, T とすると，顔の特徴点 p の 3 次元位置 (X_p, Y_p, Z_p) は以下の式で画像上の点 (x_p^i, y_p^i) に投影される．

$$k \begin{bmatrix} x_p^i \\ y_p^i \\ 1 \end{bmatrix} = A_{c_i} \left(R_{c_i}^\top \left(R \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + (T - T_{c_i}) \right) \right). \quad (1)$$

ここで A_{c_i} はカメラ i の内部パラメータ， R_{c_i} と T_{c_i} はカメラ i の回転行列，並進ベクトルを表す．頭部の位置姿勢 R と T は次式で計算される再投影誤差を最小にすることで求めることができる．

$$\{R, T\} = \arg \min_{R, T} \sum_i \sum_p f_{i,p} \left(\left(\begin{bmatrix} \hat{x}_p^i \\ \hat{y}_p^i \end{bmatrix} - \begin{bmatrix} x_p^i \\ y_p^i \end{bmatrix} \right)^\top \left(\begin{bmatrix} \hat{x}_p^i \\ \hat{y}_p^i \end{bmatrix} - \begin{bmatrix} x_p^i \\ y_p^i \end{bmatrix} \right) \right) \quad (2)$$

$$f_{i,p} = \begin{cases} 1 & (\text{feature } p \text{ is observed by camera } i.) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

3.4 虹彩の抽出

前節で得られた頭部位置・姿勢に基づいて，目領域を観測可能なカメラを選択し，選択された各カメラにおける左右の各目領域画像を抽出する（図 4）．目領域の抽出には，ベジェ曲線を用いてモデル化した瞼形状の当てはめ結果と眼球領域の境界を利用する．

抽出後，各画素ごとの輝度値に基づいて虹彩・白目のラベル付けを行う (図 5)．最後に得られた虹彩領域に対して 3 次元虹彩モデルを当てはめることで虹彩中心位置を得る．

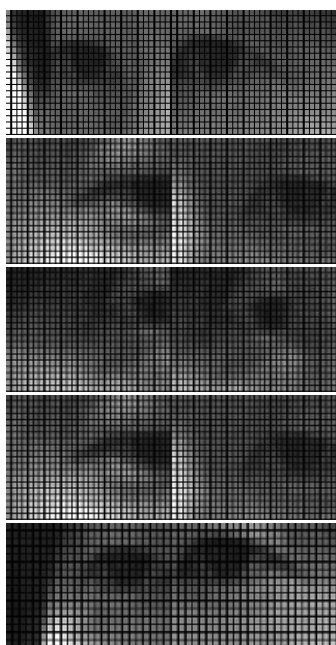


図 4 目領域画像の抽出例
Fig. 4 Examples of eye region extraction

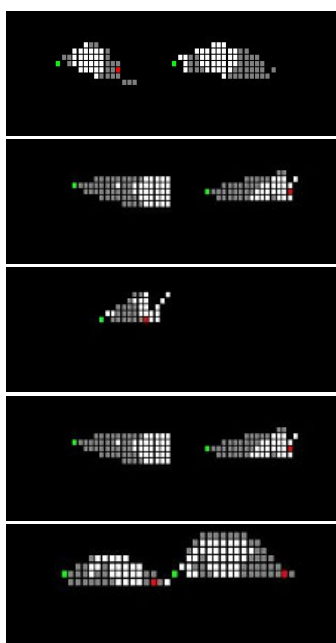


図 5 虹彩抽出結果
Fig. 5 Examples of iris extraction

3.5 視線方向の推定

提案手法では眼球の 3 次元モデルと虹彩中心を利用

して視線方向を推定する．前節までの処理で得た眼球中心位置の画像上での投影座標を $x_c = [x_c, y_c]$ ，虹彩中心の 2 次元座標を $x_{iris} = [x_{iris}, y_{iris}]$ とし，眼球領域の半径を r とすると視線方向 (カメラの光軸と視線方向との間の角度) は以下の式により推定できる．

$$\begin{aligned}\alpha &= \sin^{-1} \left(\frac{x_{iris} - x_c}{r} \right), \\ \beta &= \sin^{-1} \left(\frac{y_{iris} - y_c}{r} \right)\end{aligned}\quad (4)$$

ここで， α, β は x 軸， y 軸方向のそれぞれの視線方向である．

4. 実験

提案手法の有効性を確認するため，以下の実験を行った．図 6 実験環境はカメラ (Pointgray GRAS-50S5M-C) 2 台 (被験者からみて右方向に設置したカメラを cam1，左方向を cam2 とする) を 1.8m 間隔に配置し，1024 × 768 の解像度で撮影した (以下の実験において典型的な顔のサイズは 320 × 240 程度) ．

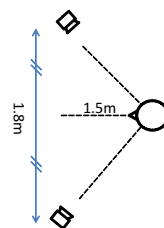


図 6 実験環境
Fig. 6 Experimental environment

まず，実装したシステムを用いて頭部姿勢の推定について評価を行った．図 7 に頭部姿勢の追跡結果の例を示す．ここでは被験者が 2 台のカメラから等距離にあって両カメラの midpoint から 1.5m 離れた位置に立ち，水平面上で左方向 (-60°) から右方向 (60°) に顔を動かした．図 7 上に見られるように提案法により連続的に頭部姿勢の追跡が行われていることがわかる．

同図下はこのときの左右の眼領域に関する各カメラ画像上での計測可否についての判定結果を示している．この結果からこのときの被験者について頭部姿勢に依らず左右の目領域が少なくともいずれか一方のカメラで取得できると判定されていることがわかる．提案法では，この判定結果に基づいて虹彩抽出に利用する画像領域を決定する．

続いて図 8 は視線推定の処理結果の例を示す．ここでは図 7 と同様に被験者は水平方向に顔向きを変化させている．ここにみられるように，2 台のカメラの観測を組み合わせることで適切に頭部姿勢および視線方

向が推定されていることがわかる．特に，一方のカメラが完全に側方からの観測になったり（同図 2 段目および最下段），2 台のカメラで共通する観測が得られない場合（同 4 段目）でも視線推定が行われていることに注意されたい．

最後に，表 1 に視線推定精度を示す．ここでは，環境中の方向既知の同一点を注視しながら顔を水平方向に変化させ，各顔向き条件についての視線推定誤差を算出した． 35° まで頭部向きを変化させた際の平均誤差は 8° 未満であった．

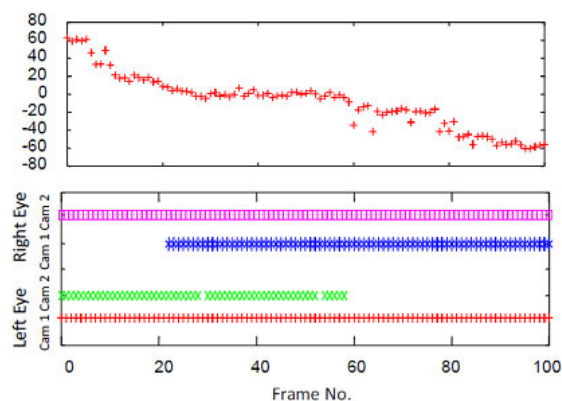


図 7 頭部姿勢追跡結果

Fig. 7 Results of head tracking

表 1 視線推定誤差（平均）

Table 1 Gaze estimation errors (average)

頭部角度 $[\circ]$	水平方向 $[\circ]$	垂直方向 $[\circ]$
0	3.3	6.2
25	7.2	6.9
35	10.1	9.5
平均誤差	6.9	7.5

5. 遠隔視線計測アプリケーション

遠隔視線計測手法を利用すると，日常環境の中で人々の興味や注意に関する情報を獲得することができ，これまで困難だった様々な様々なサービスが可能になる．

例えば米澤らは遠隔視線計測システムとロボットとを連動させることでロボットと人が視線を介してコミュニケーションできるシステムを提案した¹⁴⁾．このシステムは，ロボットがユーザと一緒に物を見る行動（共同注視）やロボットをユーザの方向へ向かせ話しかける（働きかけ）をすることによりロボットとユーザ間でコミュニケーションを誘発させるシステムである．

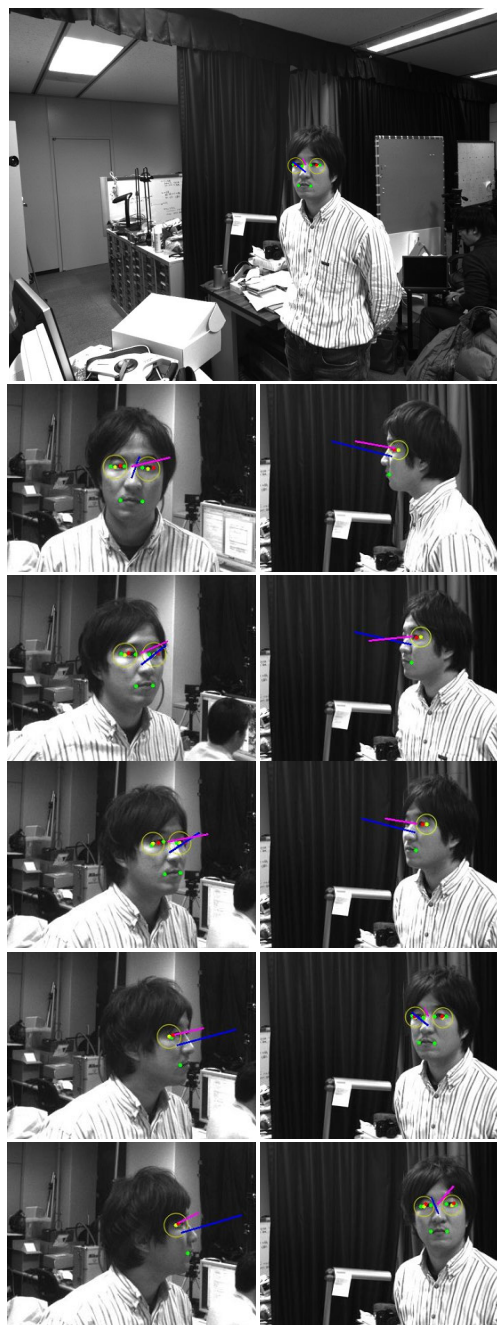


図 8 視線推定結果

Fig. 8 Results of gaze estimation

このように，視線情報を日常環境で自然なかたちで利用するためには，ユーザの頭部位置・姿勢を限定せず，できるだけ広い範囲で視線情報を計測できることが望ましい．

図 9 は，視線計測を利用した案内システムの例を示している．ここでは，案内板の前にいるユーザの興味を視線によって計測し，ユーザの興味に合わせた情報を画像や音声により案内することを想定している．しかし，2 節で述べたように従来のシステムでは特に頭

部姿勢に関する制約が大きいため、広い顔向き範囲で連続的に視線計測を行うためには、多数のカメラを利用する必要がある(同図(A))。本論文で提案した複数カメラを利用した視線計測手法では、広い顔向き範囲についてより少ないカメラで視線方向を計測できる。運転中のドライバーの視線など、大きな顔向き変化を想定しなければならない場面は多く、提案手法によってこれまで適用の難しかった分野での視線情報の利用が可能になると考えられる。

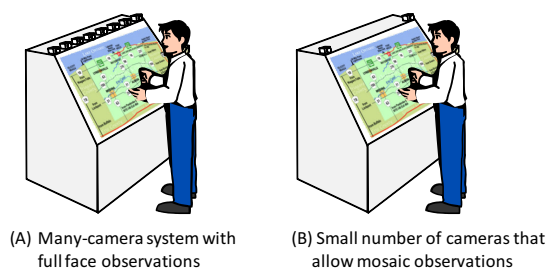


図9 提案手法による視線アプリケーションの例
Fig. 9 Example application of the proposed gaze estimation method

6. ま と め

本論文では複数のカメラの観測を組み合わせることで、従来の視線計測手法の持つ顔向きに関する制約を緩和し、広い頭部姿勢範囲について視線計測が可能な遠隔視線計測手法を提案した。提案手法では、顔面上の特徴点および左右の目領域画像の観測がそれぞれ異なるカメラで行われることを許容し、頭部の位置・姿勢および視線方向の推定を行う。そのため少数のカメラによって広い頭部姿勢範囲をカバーできる。提案手法により日常生活や公共の場における視線アプリケーションの適用範囲を大幅に拡大できると考えられる。

今後はカメラ数を増やして広域での視線計測の可用性を確認するとともに、推定精度向上のための顔、眼球モデルの更新についてもあわせて検討を進める予定である。

謝辞 本研究の一部は総務省(SCOPE)の研究委託により実施したものである。

参 考 文 献

- 1) Ohno, T., Mukawa, N., and Kawato, S. 2003. Just blink your eyes: A head-free gaze tracking system. In Proc. CHI2003, 950-951.

- 2) Morimoto, C. H., and Mimica, M. R. M. 2005. Eye gaze tracking techniques for interactive applications. Computer Vision and Image Understanding 98, 1, 4-24.
- 3) Baluja, S., and Pomerleau, D. 1994. Non-intrusive gaze tracking using artificial neural networks. Tech. Rep. CMU-CS-94-102, CMU.
- 4) Shiele, B., and Waibel, A. 1995. Gaze tracking based on face-color. In Proc. Int 'l Workshop on Automatic Face and Gesture Recognition, 344-349.
- 5) Morency, L. P., Christoudias, C., and Darrell, T. 2006. Recognizing gaze aversion gestures in embodied conversational discourse. In Proc. ICMI '06, 287-294.
- 6) Matsumoto, Y., and Zelinsky, A. 2000. An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In Proc. Int. Conf. Automatic Face and Gesture Recognition, 499-504.
- 7) 三宅哲夫, 春田誠司, 堀畑聡: 顔向きに依存しない特徴量を用いた注視判定法 電子情報通信学会論文誌・D- , 情報・システム, -パターン処理 J86-D- (12), 1737-1744, 2003-12-01
- 8) Ishikawa, T., Baker, S., Matthews, I., and Kanade, T. 2004. Passive driver gaze tracking with active appearance models. In Proc. 11th World Congress on Intelligent Transportation Systems.
- 9) 山添大丈, 内海章, 米澤朋子, 安部伸治: 単眼カメラを用いた視線推定のための三次元眼球モデルの自動キャリブレーション. 電子情報通信学会論文誌・D, 情報・システム J94-D(6), 988-1006, 2011-06-01
- 10) Taj, M., Cavallaro, A. 2011. Distributed and Decentralized Multicamera Tracking. In Proc. IEEE Signal Processing Magazine, vol.28 issue:3, 46-58
- 11) Alper, Y., Omar, J, Mubarak, S. 2006. Object tracking: A survey. Published in: journal ACM Computing Surveys(CSUR), vol.38 issue4.
- 12) Viola, P., and Jones, M. 2001. Rapid object detection using a boosted cascade of simple features. In Proc. CVPR2001, vol. 1, 511-518.
- 13) Manjunath, B. S., Chellappa, R., and Von Der Mals-Burg, C. 1992. A feature based approach to face recognition. In Proc. ICCV92, vol. 1.
- 14) Yonezawa, T., Yamazoe, H., Utsumi, A., and Abe, S. 2007. Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze-tracking. 140-145.