

MÉTODO DE CALIBRAÇÃO DE RASTREADORES DE OLHAR MÓVEIS UTILIZANDO 2 PLANOS PARA CORREÇÃO DE PARALAXE

Andrew T. N. Kurauchi*, Antonio Diaz Tula* e Carlos H. Morimoto*

*Departamento de Ciência da Computação, IME-USP, São Paulo, Brasil
e-mail: kurauchi@ime.usp.br

Resumo: As técnicas de estimação do olhar em rastreadores de olhar móveis assumem que o usuário está sempre olhando para um mesmo plano. Essa suposição é aceitável quando o olhar do usuário fica limitado a um monitor de computador, mas introduz um erro quando o rastreador é utilizado para capturar o olhar sobre objetos 3D em ambientes naturais. Um componente significativo desse erro é causado pela paralaxe entre o olho e a câmera da cena do rastreador móvel. Neste trabalho apresentamos um novo método que explora propriedades invariantes da geometria projetiva para a compensação desse erro de paralaxe com o uso de dois planos de calibração. Por meio de simulações, mostramos que o método é viável e promissor, sendo bem mais robusto a variações de profundidade do que os métodos tradicionais, apresentando erros médios até 2.5 vezes menores.

Palavras-chave: Rastreador de olhar móvel, paralaxe, geometria projetiva, calibração com dois planos.

Abstract: *Most mobile eye trackers assume that the user is always looking at the same plane. This assumption is acceptable when the user's gaze is limited to a computer screen. However it introduces an error when the eye tracker is used to collect gaze data over 3D objects in natural scenes. A significant error component is due to the parallax between the eye and the scene camera of the mobile eye tracker. In this work we present a new method based on invariant properties of projective geometry to compensate the parallax error using two calibration planes. Simulation results indicate that the method is sound and promising, being more robust to depth variations than traditional methods, with average errors up to 2.5 times smaller.*

Keywords: *Mobile eye gaze tracker, parallax, projective geometry, two planes calibration.*

Introdução

Rastreadores do olhar são utilizados em pesquisas sobre movimentos óculo-motores há várias décadas [1]. Esses dispositivos foram utilizados para auxiliar estudos sobre o comportamento do olhar durante a leitura [1], ou de indivíduos com autismo [2], entre outras aplicações [3]. Além disso, rastreadores do olhar são utilizados também como modo de interação com computadores, em particular por pessoas com deficiências físicas.

Atualmente a maioria dos rastreadores de olhar utiliza câmeras de vídeo para capturar imagens do olho.

As imagens são processadas usando algoritmos de visão computacional para a extração de características como a posição do centro da pupila (ou da íris) e seu contorno. Tradicionalmente as pesquisas sobre movimentos oculares utilizam rastreadores de olhar remotos colocados próximos a monitores de computador e não permitem grande movimentação da cabeça, sendo comum a utilização de apoios de queixo durante a execução dos experimentos. Antes do início da sessão é necessário um processo de calibração no qual o usuário deve olhar para alguns pontos conhecidos em um plano para que uma função de mapeamento entre as características do olho e a posição observada seja estimada.

Mais recentemente, com o avanço e miniaturização de computadores e câmeras de vídeo, foram desenvolvidos rastreadores de olhar móveis, possibilitando o surgimento de novas aplicações. Tipicamente, tais rastreadores utilizam um conjunto de duas ou mais câmeras montadas sobre a cabeça do usuário, uma apontando para a cena e a outra para os olhos. Esses novos dispositivos permitem a realização de experimentos mais realistas, onde os dados podem ser obtidos durante a execução de atividades no ambiente real, como por exemplo, durante a prática esportiva.

Apesar dos avanços no hardware, a maioria dos sistemas móveis ainda utiliza as mesmas hipóteses de rastreadores remotos, ou seja, que o usuário está olhando para pontos em um plano a uma distância fixa. Isso causa um erro de paralaxe devido à distância entre os olhos e a câmera da cena quando a distância ao plano observado se altera. Dessa forma um modelo geométrico mais complexo, que considere a distância entre o olho e a câmera da cena, deve ser considerado para que os rastreadores de olhar móveis possam ser utilizados em ambientes mais naturais.

Um estudo sobre a distância ideal de calibração para minimizar o erro de paralaxe em uma determinada região foi realizado por Mardanbegi e Hansen [4]. Além disso, algumas técnicas para estimação do olhar com compensação do erro de paralaxe foram propostas. Entretanto elas utilizam sistemas ópticos complexos [5], projetores infravermelhos de alta potência com uma iluminação ativa codificada [6], ou câmeras calibradas [7], dificultando a sua utilização em ambientes não controlados.

A contribuição desse trabalho é a introdução de um novo método de estimação baseado em propriedades

invariantes da geometria projetiva [8] para a compensação do erro de paralaxe, que permitirá a construção de rastreadores móveis com maior precisão e acurácia na medição dos movimentos oculares usando câmeras não calibradas. A seguir descrevemos a técnica e resultados de simulação que demonstram a sua viabilidade e robustez.

Rastreador de olhar móvel

A figura 1 mostra um rastreador de olhar móvel, composto por uma câmera da cena C_s apontando para frente do usuário e uma câmera do olho C_e , que aponta para o olho do usuário. Utilizaremos C_c , P_c e Π_c para denominar o centro de projeção, a matriz de projeção e o plano da imagem da câmera C , respectivamente.

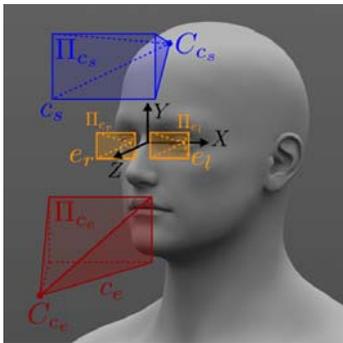


Figura 1: A câmera da cena (C_s) aponta para frente, capturando a cena observada pelo usuário. A câmera do olho (C_e) aponta para os olhos, que são modelados como câmeras *pinhole* (e_l e e_r) com os centros de projeção (C_{e_l} e C_{e_r}) localizados no centro das córneas. O sistema de coordenadas é definido pela origem no ponto médio entre C_{e_l} e C_{e_r} , eixo X apontando para a esquerda, eixo Y pra cima e Z para a frente do usuário.

O processo de rastreamento do olhar deve capturar imagens dos olhos, estimar suas orientações e calcular o ponto na cena sendo observado. Para permitir a observação dos pontos observados é utilizada uma câmera que captura a cena. Assim, a estimação do olhar em rastreadores móveis pode ser feita mapeando características do olho diretamente na câmera da cena.

Uma maneira simples e elegante de se calcular esse mapeamento é por meio de homografias [9]. Para isso, além das câmeras do olho e da cena, os olhos também são modelados como câmeras *pinhole* e_l e e_r (para o olho esquerdo e direito, respectivamente). Para simplificar o modelamento geométrico, vamos assumir que o plano de formação da imagem do olho é paralelo ao plano XY e tangencia a córnea. O centro de projeção de cada olho se encontra no centro da córnea. Utilizaremos o índice j para se referir a qualquer dos olhos (l ou r).

Como as câmeras devem ser montadas sobre uma armação, vamos considerar que elas permanecem fixas com relação aos olhos. Nesse caso, utilizando propriedades da geometria projetiva [8], sabemos que a

projeção perspectiva entre 2 planos e definida por uma homografia H_a^b que mapeia pontos do plano Π_a para Π_b . Como a composição de homografias também é uma homografia, podemos calcular uma transformação de pontos no plano da imagem de C_e para pontos no plano da imagem de C_s , ou seja, existe uma transformação do plano de imagem do olho e_j para a câmera C_e dada por $H_{e_j}^{C_{e_j}}$, e outra transformação de C_e para a câmera da cena C_s dada por $H_{C_{e_j}}^{C_s}$, cuja composição $H_{e_j}^{C_{e_j}} H_{C_{e_j}}^{C_s}$ equivale a $H_{e_j}^{C_s}$. Assumindo que o centro da pupila p_j esteja no plano de imagem do olho, então p_j pode ser mapeado diretamente para um ponto g em C_s pela fórmula $g = H_{e_j}^{C_s} p_j$. A homografia pode ser calculada por um processo de calibração, onde o usuário deve fixar o olhar sobre ao menos 4 pontos de referência em um plano de calibração Π_1 [8].

Esse método é bastante atraente pela sua simplicidade e uso de câmeras não calibradas. No entanto, a acurácia desse método para pontos fora do plano utilizado para calibração prejudica o seu uso em ambientes naturais. Um componente significativo desse erro é devido à paralaxe entre o olho e a câmera da cena, como descrito a seguir.

Erro de paralaxe

O erro de paralaxe é devido à distância entre a câmera da cena e o olho. Considere o caso em que o usuário esteja olhando para um ponto $q \notin \Pi_1$ no espaço, ou seja, fora do plano de calibração. Os pontos q e C_{e_j} definem a direção do olhar v_q^j . Sejam q^{Π_1} a intersecção de v_q^j com Π_1 e p_q^j o centro da pupila associada a v_q^j .

Se a câmera da cena pudesse ser colocada sobre o olho (caso ideal onde C_{c_s} é igual a C_{e_j}) então q , $P_{C_s}(q)$ (a projeção de q no plano da imagem da câmera C_s), q^{Π_1} , $P_{C_s}(q^{\Pi_1})$, $P_{C_e}(q)$, e C_{e_j} seriam todos colineares e $P_{C_s}(q) = P_{C_s}(q^{\Pi_1})$. Nesse caso temos $H_1^j \cdot P_{C_e}(p_q^j) = P_{C_s}(q^{\Pi_1}) = P_{C_s}(q)$.

Assim H_1^j seria suficiente para estimar a projeção de qualquer ponto observado no espaço em C_s . Todavia, utilizando somente câmeras de vídeo tradicionais, C_{c_s} não será igual à C_{e_j} . Devemos então analisar o caso onde $C_{c_s} \neq C_{e_j}$.

Seja $\delta_q := q - q^{\Pi_1}$ de tal forma que $P_{C_s}(q^{\Pi_1}) = P_{C_s}(q - \delta_q)$, e portanto $P_{C_s}(q) = P_{C_s}(q^{\Pi_1}) + \epsilon_q$, onde $\epsilon_q := P_{C_s}(\delta_q)$, que em muitos casos é não nulo. Em outras palavras, se $C_{c_s} \neq C_{e_j}$ e o usuário está olhando para $q \notin \Pi_1$ então a aplicação da homografia H_1^j sobre a imagem do centro da pupila associada $P_{C_e}(p_q^j)$ resultará em um ponto em Π_{C_s} deslocado de um erro ϵ_q . Note que se $q \in \Pi_1$ então $\delta_q = \mathbf{0}$, então ϵ_q é definido para todos os pontos q no espaço 3D. Definimos o erro de paralaxe causado pela distância entre C_{c_s} e C_{e_j} como esse deslocamento de ϵ_q em Π_{C_s} (Figura 2).

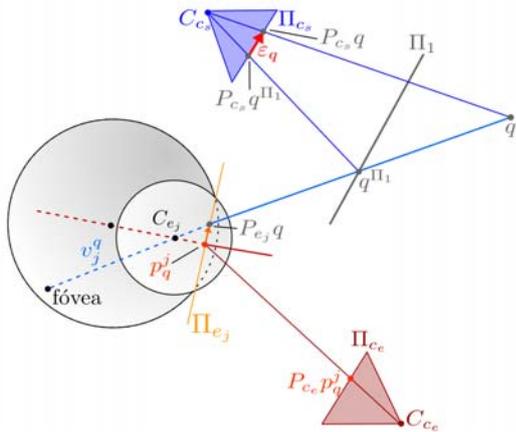


Figura 2: O erro de paralaxe pode ser definido como a distância ϵ_q entre as projeções da posição estimada do olhar ($P_{C_s}(q^{III})$) e da posição real do olhar ($P_{C_s}(q)$).

Método de calibração com dois planos

A Figura 3 ilustra o princípio de funcionamento da nossa técnica, que explora o fato de que pontos colineares continuam colineares após passarem por projeção perspectiva [8].

Seja g o ponto observado no espaço em um dado momento. Para estimar $P_{C_s}(g)$ utilizaremos a noção do movimento de vergência, a rotação dos olhos em direções opostas para manter a projeção do objeto observado na fóvea em diferentes distâncias. Devido a esse movimento sabemos que a menor distância entre as direções do olhar v_g^r e v_g^l ocorrerá próxima de g . Conseqüentemente a intersecção das linhas $P_{C_s}(v_g^j)$ (projeções das direções do olhar v_g^r e v_g^l em C_s) também será próxima de $P_{C_s}(g)$. Assim, $P_{C_s}(g)$ pode ser estimado como a intersecção de $P_{C_s}(v_g^l)$ e $P_{C_s}(v_g^r)$.

Para se estimar as linhas $P_{C_s}(v_g^j)$, considere as homografias H^j e H^j resultantes da calibração com dois planos distintos Π_1 e Π_2 . Sejam p_g^l e p_g^r os centros das pupilas associados às direções do olhar v_g^r e v_g^l . Os pontos $g_{1,j} = H^j \cdot P_{C_e}(p_g^j)$ e $g_{2,j} = H^j \cdot P_{C_e}(p_g^j)$ correspondem às projeções das intersecções de v_g^j com Π_1 e Π_2 na imagem da cena e estão contidos na reta definida pela direção do olhar v_g^j . Assim, cada pupila é mapeada 2 vezes (usando Π_1 e Π_2) em C_s , definindo uma linha correspondente à direção do olhar daquela pupila, e a posição de $P_{C_s}(g)$ é definida pela intersecção das duas retas, cada uma definida pela posição dos centros das pupilas detectadas nas imagens das câmeras dos olhos.

Resultados

Para a validação da técnica foi utilizado o simulador de rastreadores de olhar desenvolvido por Böhme et al. [10], que utiliza valores médios encontrados na literatura para as dimensões das estruturas dos olhos e coeficientes de refração.

Configuração – Para a calibração foram utilizados os planos Π_1 e Π_2 a 0.65m e 1.65m de distância dos

olhos. Em cada plano são utilizados 4 pontos, para o cálculo das homografias H^j e H^j , nos cantos de um retângulo de 0.40 m de largura por 0.30 m de altura centralizado na altura dos olhos.

As duas câmeras se encontram 0.05 m a frente dos olhos. A câmera da cena aponta para o centro do monitor em Π_1 e a câmera do olho aponta para o centro dos olhos.

Para comparar o desempenho quanto à robustez a mudanças de profundidade, cada técnica foi aplicada a cada 0.10 m a planos entre 0.65 m e 5.65 m. Para cada plano foi calculada a média e desvio padrão dos erros de 49 pontos apresentados, distribuídos uniformemente em uma grade de 7 por 7 pontos.

Técnicas utilizadas – O método proposto foi comparado com a calibração tradicional com um plano utilizando uma homografia. Assim, para cada plano de teste foram utilizados os resultados da técnica proposta ($2p$), o ponto médio entre os pontos estimados por H^j para os dois olhos ($1n$) e o ponto médio entre os pontos estimados por H^j para os dois olhos ($1f$).

Cálculo do erro – Cada ponto estimado na imagem da cena foi projetado no plano atual de testes para obter a posição estimada do olhar. O erro, em graus, é definido como o ângulo formado pelas posições estimada e real do olhar com o centro do olho.

Os resultados obtidos podem ser observados na Figura 4. As médias dos erros de todas as distâncias para as técnicas $2p$, $1n$ e $1f$ foram, respectivamente, 2.14, 5.40 e 3.01 graus. Assim, o erro médio obtido pela calibração com dois planos foi 2.52 vezes menor do que o de $1n$ e 1.41 vezes menor do que o de $1f$.

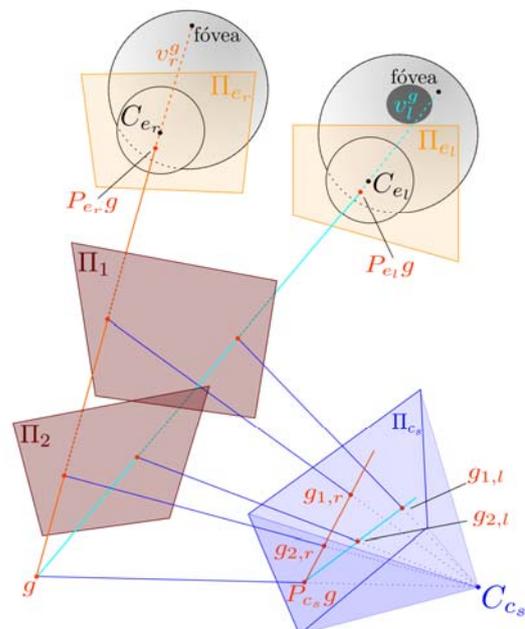


Figura 3: A calibração com dois planos consiste em utilizar 4 homografias (2 para cada olho) para calcular a intersecção das projeções das retas definidas pelas direções do olhar na imagem da cena.

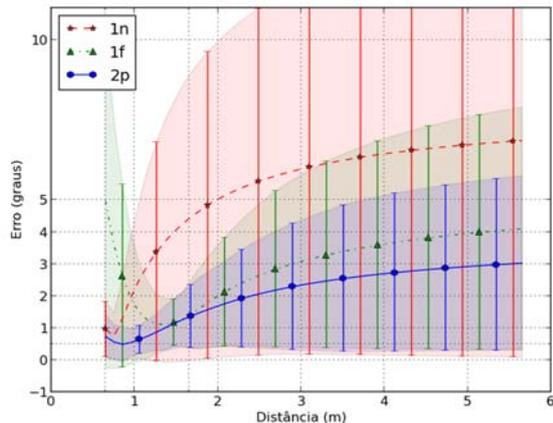


Figura 4: Erros médios obtidos na simulação para a calibração com dois planos ($2p$), a calibração usando somente H_1 ($1n$) e somente H_2 ($1f$). São apresentados intervalos de um desvio padrão para cima e para baixo. Os dados apresentados foram extraídos do simulador de Böhme et al. [10].

Discussão

Os erros médios até 2.5 vezes menores indicam uma maior robustez do método proposto a variações de profundidade do que os métodos tradicionais.

É possível observar um erro maior próximo às distâncias de calibração na técnica proposta. Uma possível explicação é que mesmo no plano de calibração os valores estimados por H_1^l e H_2^l não são coincidentes, mas são bastante próximos. Assim, o ponto médio entre as estimações para os dois olhos é uma boa aproximação. Entretanto, devido à diferença entre os pontos estimados, ocorre um deslocamento na intersecção das retas calculadas, apresentando um erro maior. O uso de 3 ou mais planos de calibração poderia ajudar a tornar o sistema ainda mais robusto, porém tornaria o processo de calibração mais complexo.

Conclusões

Neste trabalho foi apresentada uma nova técnica de calibração de rastreadores de olha móveis utilizando dois planos. Devido ao fato da projeção perspectiva preservar a colinearidade dos pontos [8] é possível estimar a projeção das direções do olhar a partir da projeção das intersecções da direção do olhar com dois planos distintos. Essas intersecções são obtidas através da aplicação de duas morfografias sobre a imagem do centro da pupila. Devido ao movimento de vergência dos olhos é possível estimar a posição observada como a intersecção das duas direções.

O resultado de uma simulação mostra que a calibração com dois planos é viável e promissora. A sua maior robustez a variações de profundidade é evidenciada por erros médios até 2.5 vezes menores do que dos métodos tradicionais.

Estamos desenvolvendo um protótipo utilizando 2 câmeras, uma para a cena e outra para ambos os olhos, montadas em um capacete para a realização de testes reais.

Agradecimentos

Agradecemos o apoio financeiro da FAPESP, processos nº 2011/00267-1, 2012/04426-0 e 2013/06791-0, Fundação de Amparo à Pesquisa do Estado de São Paulo.

Referências

- [1] Rayner K. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*. 1998 Nov; 124(3):372–422.
- [2] Boraston Z, Blakemore SJ. The application of eye-tracking technology in the study of autism. *The Journal of Physiology*. 2007 Jun; 581(Pt 3):893–8.
- [3] Duchowski AT. A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*. 2002 Nov; 34(4):455–70.
- [4] Mardanbegi D, Hansen DW. Parallax error in the monocular head-mounted eye trackers. *Proceedings of ACM Conference on Ubiquitous Computing*. 2012; 689-94.
- [5] Mori H, Sumiya E, Mashita T, Kiyokawa K, Takemura H. A wide-view parallax-free eye-mark recorder with a hyperboloidal half-silvered mirror and appearance-based gaze estimation. *IEE Transactions on Visualization and Computer Graphics*. 2011; 17(7):900-12.
- [6] Nakazawa A, Nitschke C. Point of gaze estimation through corneal surface reflection in an active illumination environment. *Proceedings of the 12th European Conference on Computer Vision*. 2012; Part II:159-72.
- [7] Pirri F, Rudi A. A general method for the point of regard estimation in 3D space. *IEEE Conference on Computer Vision and Pattern Recognition*. 2011; 921-8.
- [8] Hartley R, Zisserman A. *Multiple View Geometry in Computer Vision*. 2^a ed. Cambridge University Press; 2004.
- [9] Li D, Winfield D, Parkhurst DJ. Starburst: A Hybrid Algorithm for Video-based Eye Tracking Combining Feature-based and Model-based Approaches. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition – Workshops*. 2005; 79.
- [10] Böhme M, Dorr M, Graw M, Martinetz T, Barth E. A Software Framework for Simulating Eye Trackers. *Proceedings of the 2008 Symposium on Eye Tracking Research and Applications*. 2008; 251-8.
- [11] Morimoto CH, Mimica MRM. Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding*. 2005; 98(1):4-24.