

Detecção da Direção do Olhar via *Webcam*

Rafael Gomes Monteiro¹, Cabral Lima², Antonio Carlos Gay Thomé³

Programa de Pós-Graduação em Informática (PPGI)

Universidade Federal do Rio de Janeiro (UFRJ)

Rio de Janeiro, Brasil

E-Mail: ¹rafaelgomesmonteiro@gmail.com, ²cabrallima@ufrj.br, ³thome@nce.ufrj.br

Resumo—Dentre as formas de Interação Humano-Computador descritas na literatura existem os sistemas de rastreamento do olhar, que consistem em estimar o ponto na tela para onde o usuário está focando a visão. Neste trabalho¹ foi realizado um estudo sobre os sistemas de rastreamento do olhar existentes na literatura e foi desenvolvido um protótipo funcional que apresentou resultados bastante significativos, conseguindo alcançar cerca de 5.6° de precisão nos experimentos realizados. As principais contribuições deste trabalho concernem à simplicidade e ao baixo custo de implementação porque é utilizada uma *webcam* comum para capturar as imagens, e o uso de iluminação ambiente, enquanto que a maioria dos sistemas utiliza câmeras de alta resolução e iluminação infravermelha, o que aumenta o custo.

Abstract—Amongst the several existing ways of human-computer interaction described in the literature there are the eye tracking systems, which estimate the point on the screen where the user is looking at. In this paper we did a study on the gaze tracking systems available in the literature and developed a working prototype that showed significant results, managing to reach about 5.6 degrees of precision in the experiments. The major contributions of this study are in the usage of a common webcam to capture images and the use of ambient light, as most systems use high-resolution cameras and infrared lighting, which increases their costs.

Keywords—detecção do olhar; processamento de imagens; webcam;

I. INTRODUÇÃO

O rastreamento do olhar consiste em tentar estimar o ponto na tela para onde o usuário está olhando, com base na captura de imagens dos seus olhos. Existem vários tipos de sistemas que realizam essa tarefa [2]. Em alguns sistemas o usuário utiliza um capacete que possui câmeras direcionadas para os seus olhos. Sensores estimam a posição da cabeça, enquanto a câmera fornece imagens em alta resolução dos olhos, para estimar a direção do olhar. Combinando essas informações e a posição da tela do computador é possível estimar o ponto para onde o usuário está olhando. Apesar de serem bastante precisos, há o problema desses sistemas serem invasivos, devido ao hardware que deve ser acoplado no usuário.

Existem também sistemas não-invasivos, os quais possuem tipicamente uma ou mais câmeras de vídeo posicionadas à frente do usuário, apontadas na sua direção, capturando imagens que são processadas por um *software* que tenta estimar o olhar do usuário. Apesar de geralmente serem menos precisos, possuem um custo mais baixo e são mais acessíveis [2].

Contribuições: O principal objetivo neste trabalho foi desenvolver um sistema de detecção da direção do olhar com base em imagens capturadas a partir de uma câmera de vídeo. Complementando o foco principal, estabeleceu-se como objetivo específico que o sistema possuísse baixo custo, sendo capaz de operar com uma *webcam* comum, tornando-o mais próximo de ser acessível para o público em geral.

A principal motivação para o desenvolvimento deste trabalho está relacionada às aplicações do produto gerado, pois espera-se que ele contribua para aumentar a acessibilidade de pessoas portadoras de algum tipo de deficiência motora que as impeçam de acessar o computador via teclado e mouse.

A. Trabalhos relacionados

Segundo Hansen e Ji [2], que recentemente publicaram um levantamento sobre as pesquisas que vêm sendo realizadas na área, duas subáreas concentram o esforço e interesse dos pesquisadores: a localização do olho na imagem e a estimativa da direção do olhar.

A localização do olho na imagem visa encontrar, com precisão, a posição dos olhos em cada quadro obtido pela câmera. Isso é feito através de técnicas de processamento de imagens e reconhecimento de padrões. Algumas abordagens utilizam modelos baseados em formas fixas ou deformáveis [3, 4], construídos com base em descritores (*features*), informações sobre bordas ou reposta de filtros. Outras utilizam modelos baseados na aparência, através de casamento de padrões (*template matching*), utilizando medidas de similaridade [5, 6]. Há também os métodos híbridos, que combinam técnicas utilizadas nas duas outras abordagens [7, 8, 9].

A estimativa da direção do olhar visa tentar determinar o ponto para onde o usuário está olhando na tela. Geralmente isso é feito através de um processo de calibração, onde o usuário olha para alguns pontos pré-definidos, que são armazenados pelo sistema e interpolados para gerar a informação da direção do olhar nos quadros sucessivos de vídeo [10, 11].

II. FUNDAMENTAÇÃO TEÓRICA

Neste trabalho optou-se por utilizar técnicas que necessitem de pouco processamento, sem muitas exigências de *hardware*.

Na localização do olho, foram utilizadas técnicas baseadas em modelos de formas fixas, pois utilizam poucas variáveis livres para casar os modelos com as formas existentes na imagem [2]. Já na estimativa do olhar, foi utilizado um modelo neural para construir uma função de mapeamento entre o deslocamento dos olhos e as coordenadas de tela [12].

¹Master's thesis [1]

III. IMPLEMENTAÇÃO

O sistema foi desenvolvido em módulos, que atuam de forma sequencial, conforme ilustra a Fig. 1.

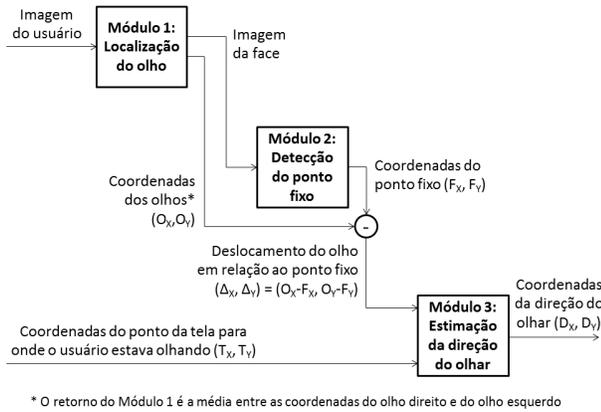


Fig. 1. Visão geral dos módulos da abordagem proposta

O Módulo 1 localiza os olhos do usuário, mais especificamente as coordenadas do centro das íris, retornando o ponto médio, denotado por (O_X, O_Y) . O Módulo 1 também é responsável por localizar a face do usuário que servirá como entrada para o Módulo 2. Este faz a detecção do ponto fixo, cujas coordenadas são representadas por (F_X, F_Y) . Ponto fixo é uma referência que permanece estacionária entre diferentes quadros de vídeo, a partir do qual é calculado o deslocamento dos olhos da seguinte forma: $(\Delta_X, \Delta_Y) = (O_X - F_X, O_Y - F_Y)$. O Módulo 3 utiliza essa informação de deslocamento para estimar a direção do olhar. Para isso, ele utiliza também as coordenadas do ponto na tela para onde o usuário estava olhando, denotadas por (T_X, T_Y) .

A. Módulo 1: Localização do olho na imagem

Inicialmente foi detectada a face do usuário, utilizando o método proposto por Viola and Jones [13]. Em seguida, localiza-se os olhos. Para reduzir o processamento, delimitou-se duas áreas de busca através de regiões na face, ilustradas na Fig. 2. Os percentuais foram obtidos de forma empírica.

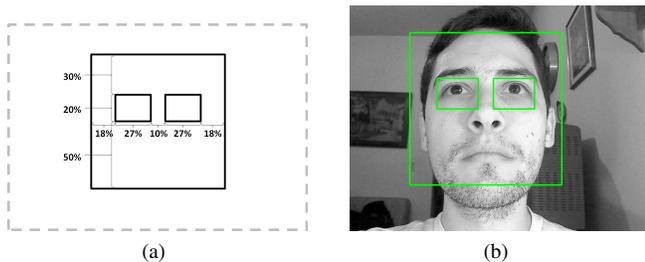


Fig. 2. (a) Sub-regiões que contém os olhos e (b) exemplo numa imagem

1) *Localização da íris:* a localização da íris na imagem consiste em localizar com precisão as coordenadas da íris. Para normalizar a luminosidade, foi realizada a expansão do

histograma, destacando a íris do restante da imagem para facilitar sua localização. Para reduzir o ruído, foi aplicado o filtro bilateral [14]. Esse filtro suaviza uma imagem preservando suas bordas, o que é útil para a localização da íris. A aplicação desse filtro na imagem do olho é ilustrada na Fig. 3.



Fig. 3. (a) Imagem do olho e (b) resultado da aplicação do filtro bilateral

Detecção da íris: O primeiro passo para esse processo é a detecção das bordas da imagem. Foram testados os filtros de Sobel [15] e Canny [16]. A Fig. 4 mostra o resultado da aplicação de ambos os filtros na imagem do olho. Nos experimentos realizados, optou-se por utilizar o filtro de Canny.



Fig. 4. Detecção de bordas pelos filtros de (a) Sobel e (b) Canny

Em seguida, é aplicada a Transformada Circular de Hough [17, 18]. O resultado da detecção é exibido na Fig. 5b através de um círculo de raio $r = 20$ pixels centrado nas coordenadas do pico do espaço de Hough, indicado pela seta na Fig. 5a.

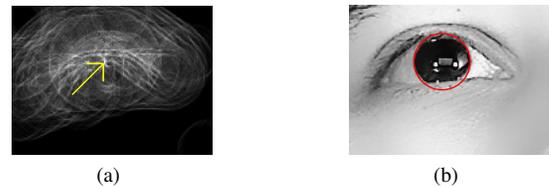


Fig. 5. (a) Espaço de Hough gerado com raio $r = 20$ pixels e (b) resultado da detecção na imagem original

O raio da íris varia de pessoa para pessoa ou por outros fatores, como a distância da câmera. Conforme sugerido em [19], foi definida uma faixa de raios possíveis em torno de um raio médio, definido como 2% da largura da face. Esta faixa é definida no intervalo de -20% a $+20\%$, com salto de 5%, gerando nove variações. Para cada variação, foi aplicada a transformada de Hough para detectar os círculos. Criou-se uma métrica para pontuar os círculos detectados para escolher o que melhor representa a íris. Como esta é geralmente mais escura do que o seu redor, foi criada uma máscara branca com um círculo inscrito em preto, para cada variação de raio. Sua aplicação é feita pelo cálculo da diferença absoluta entre

seus píxeis e os píxeis da região candidata à íris. Escolhe-se a região de menor diferença média como a íris detectada.

A Fig. 6 ilustra a aplicação desta validação. São exibidas as nove regiões candidatas e a pontuação atribuída de acordo com a máscara referente àquela região. No exemplo, o melhor raio seria $r = 21$, com pontuação igual a 0.20408.

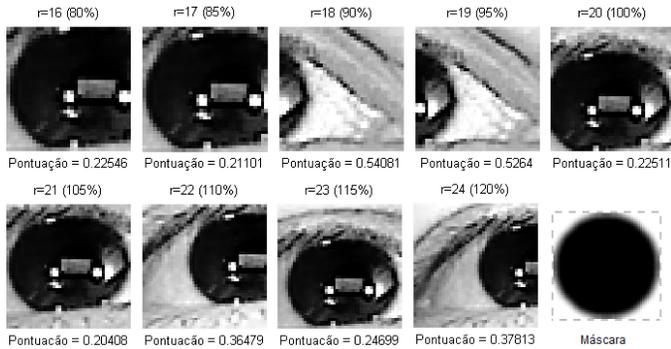


Fig. 6. Regiões candidatas e respectiva pontuação obtida pelo uso da máscara

Refinamento e estimação das coordenadas: em certos casos, as íris não são detectadas com muita precisão, principalmente quando o usuário está olhando para os lados fazendo com que a íris assuma uma forma elíptica. Foi criado um refinamento baseado numa máscara, ilustrada na Fig. 7. Os píxeis cinzas valem zero, os brancos possuem valores positivos e os pretos, negativos, de forma que a soma seja igual a zero.



Fig. 7. Máscara utilizada para o refinamento da detecção da íris

Tomando como base o melhor raio detectado anteriormente, é definida uma nova faixa de raios variando de -20% a $+20\%$ do raio, com um salto de 10% . Para cada raio, é feita a convolução da máscara com uma região ao redor do centro da íris, escolhendo-se a região com o maior pico na convolução. Isso é ilustrado na Fig. 8.

Na primeira linha estão as cinco regiões candidatas, recortadas de acordo com os raios da faixa de raios. Na segunda linha estão os resultados das convoluções das máscaras com cada imagem. As setas indicam o pico da convolução. Na terceira linha é apresentada a detecção refinada. No caso, foi selecionada a quarta coluna, de raio $r = 18.7$ píxeis e pontuação igual a 0.46048.

B. Módulo 2: Detecção do ponto fixo

Um ponto fixo ideal deve ser:

- 1) Universal: todos os usuários devem possuí-lo;
- 2) Imutável: deve manter sua forma em todos os quadros de vídeo, para que possa ser rastreado;
- 3) Estático: sua posição deve permanecer fixa com relação aos componentes da face (principalmente os olhos).

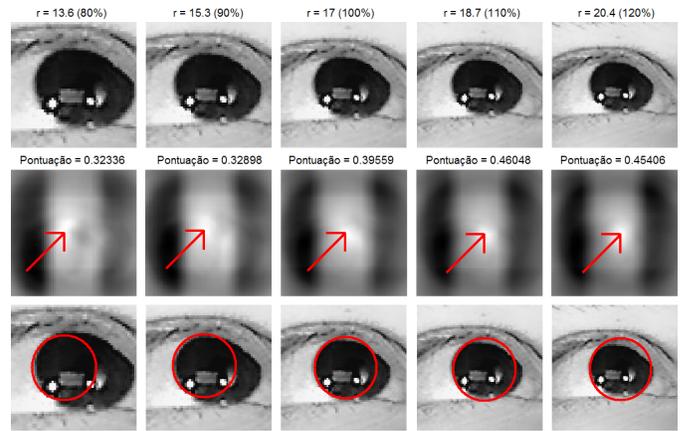


Fig. 8. Refinamento da detecção da íris

Foram feitos experimentos com três pontos fixos: nariz, sobrancelhas e uma marcação artificial posicionada na testa do usuário. A detecção do nariz foi feita de forma similar à redução da área de busca (como mostra a seção III-A), através de uma região fixa. Foi definida uma extensão das margens superior e inferior da região fixa para se fazer uma busca pela melhor região. A extensão abrange 25% da sua altura.

Essa região é varrida de cima para baixo, considerando todas as sub-regiões possíveis com as mesmas dimensões da região fixa original, com saltos de um píxel. Para cada sub-região é calculado o desvio padrão dos píxeis nela contidos. É selecionada a região com o maior desvio padrão. A Fig. 9 ilustra a detecção do nariz. As linhas na horizontal delimitam os limites superior e inferior da região de busca.

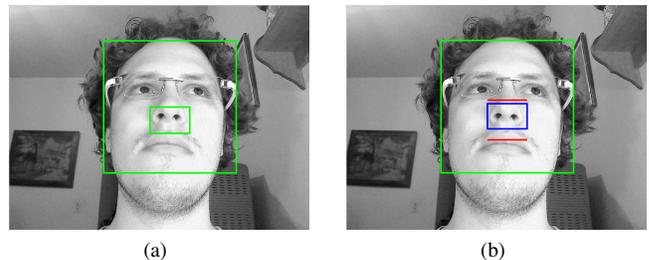


Fig. 9. Enquadramento (a) inicial e (b) ajustado do nariz

Para o rastreamento, foi definida uma área de busca maior que a área original onde o ponto fixo foi detectado. É feita uma busca nessa região de forma a tentar encontrar a imagem do ponto fixo, salva anteriormente. Foi utilizada a técnica de *Phase Correlation* [20], buscando a imagem salva na detecção.

C. Módulo 3: Estimação da direção do olhar

A estimação da direção do olhar é dividida em duas fases: calibração e estimação.

1) *Processo de calibração:* o processo consiste em exibir uma série de pontos para o usuário, que olha para os mesmos enquanto o sistema captura sua imagem através da *webcam*.

Para cada ponto, foram capturadas cinco imagens para medir a estabilidade do sistema quando o usuário olha fixamente para o mesmo ponto. Combinando a informação do deslocamento dos olhos ($\Delta x, \Delta y$) e do ponto para onde o usuário estava olhando (T_x, T_y), foi treinada uma Rede Neural Artificial (RNA). A Fig. 10 ilustra um exemplo de dados de treinamento.

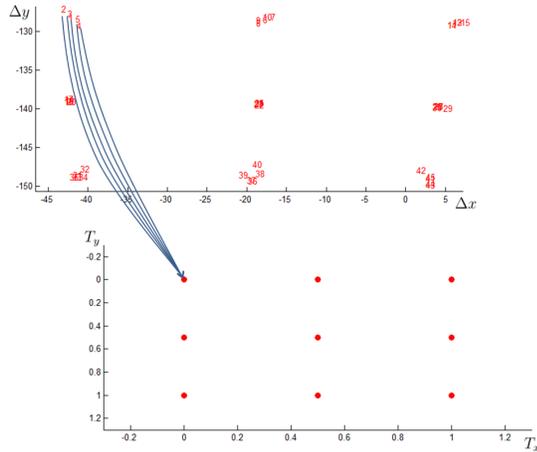


Fig. 10. Ilustração dos dados de treinamento da RNA, que deverá mapear as entradas (coordenadas do deslocamento do olho ($\Delta x, \Delta y$)) para as saídas (coordenadas de tela (T_x, T_y))

2) *Estimação*: após a calibração, a estimação do olhar pode ser feita passando as coordenadas do deslocamento do olho ($\Delta x, \Delta y$) dos dados de teste para a RNA, que irá retornar as coordenadas estimadas da direção do olhar (D_x, D_y).

IV. EXPERIMENTOS

Na calibração, foram exibidos nove pontos dispostos numa grade 3x3. No teste da estimação foram exibidos 25 pontos (5x5). Esses pontos são ilustrados na Fig. 11.

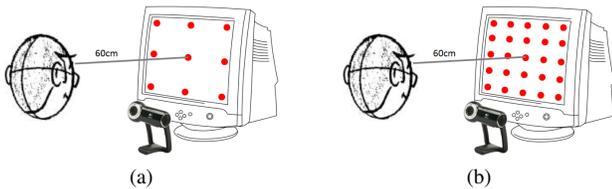


Fig. 11. Pontos utilizados nas fases de (a) calibração e (b) teste da estimação

Para os experimentos, foi criada uma base de imagens num ambiente controlado. Foram capturadas imagens de cinco usuários distintos. O usuário permanecia a uma distância de 60cm de um monitor com tela de 24". A *webcam* utilizada foi uma Logitech QuickCam Pro 9000, capturando imagens na resolução de 1600x1200 píxeis. Para cada usuário, foram feitas duas capturas: uma com o marcador artificial e uma sem, para testar o uso do nariz e das sobrancelhas como pontos fixos. Para evitar que os usuários movimentassem a cabeça durante o experimento, foi improvisado um suporte para a cabeça. A câmera foi fixada no suporte, para que sua posição em relação

ao usuário se mantivesse fixa. A iluminação foi feita com duas luminárias, posicionadas de cada lado do monitor. Uma visão geral do ambiente é ilustrada na Fig. 12.

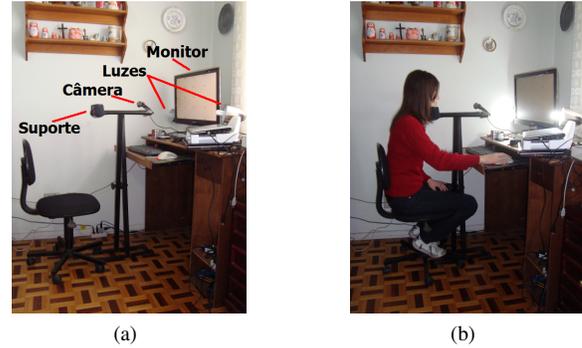


Fig. 12. (a) Ambiente de captura e (b) usuário no ambiente

V. RESULTADOS E ANÁLISES

A seguir são apresentados e discutidos os resultados obtidos em cada módulo.

A. Localização do olho na imagem

Para medir a acurácia da localização da íris, foi utilizada a métrica definida por Jesorsky et al. [21], por ser invariante ao tamanho da face. Ela é calculada da seguinte forma:

- 1) As íris são detectadas, gerando dois pares de coordenadas ($OEsq_x, OEsq_y$) e ($ODir_x, ODir_y$);
- 2) Calcula-se a distância euclidiana entre cada par de coordenadas e os pares correspondentes à marcação real (denominados ($MEsq_x, MEsq_y$) e ($MDir_x, MDir_y$)), marcados manualmente, gerando duas distâncias;
- 3) A maior distância é dividida pela distância entre os dois olhos, utilizando a marcação manual. Essa distância é chamada de distância relativa dos olhos.

Para fins de comparação, foi feita uma análise dos resultados da localização da íris com e sem o refinamento, apresentado na seção III-A1. Os resultados são apresentados na tabela I.

TABELA I
RESULTADOS DA LOCALIZAÇÃO DA ÍRIS. A LINHA "TOTAL" CORRESPONDE À MÉDIA DE CADA COLUNA, EXCETO NA COLUNA "QUANTIDADE DE IMAGENS", ONDE CORRESPONDE À SOMA.

Usuário	Quant. de imagens	Erro sem o refinamento	Erro com o refinamento	Diferença absoluta
1	340	0.0153	0.0150	0.0003
2	340	0.0137	0.0203	-0.0066
3	340	0.0163	0.0162	0.0001
4	340	0.0167	0.0196	-0.0029
5	340	0.0116	0.0178	-0.0062
Total	1700	0.0147	0.0178	-0.0031

Aparentemente não houve diferença significativa. Porém, observou-se que o refinamento torna as coordenadas mais estáveis. Isso ocorre porque a convolução é suavizada, enquanto que a Transformada de Hough procura fazer um

casamento exato com as bordas da imagem. Em certos casos, isso poderia gerar ambiguidade, conforme ilustra a Fig. 13, o que pode fazer com que as possibilidades de detecção sejam escolhidas alternadamente, daí a instabilidade na localização da íris. O refinamento faz com que a estimação do olhar possua uma acurácia maior, como será visto na seção V-C.

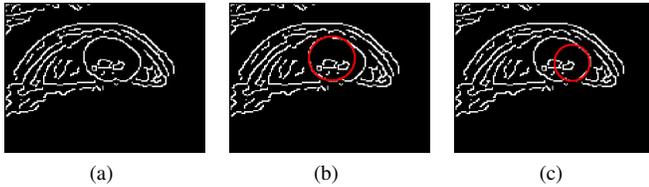


Fig. 13. Confusão que poderia ser gerada pela Transformada de Hough: (a) Imagem de bordas e (b)(c) possibilidades de detecção da íris

B. Detecção do ponto fixo

A exibição de resultados desse módulo será feita graficamente, visto que a base de imagens não possui informação sobre a posição do ponto fixo na imagem. Isso seria inviável de ser realizado porque o algoritmo seleciona esse ponto buscando a área com maior desvio padrão dos píxeis, e uma seleção manual seria feita de forma empírica.

A seguir são apresentados resultados gráficos da detecção e do rastreamento dos pontos fixos. A Fig. 14 ilustra alguns resultados. O retângulo tracejado indica a seleção do ponto fixo no primeiro quadro de vídeo e os demais retângulos indicam o seu rastreamento nos demais quadros de vídeo.

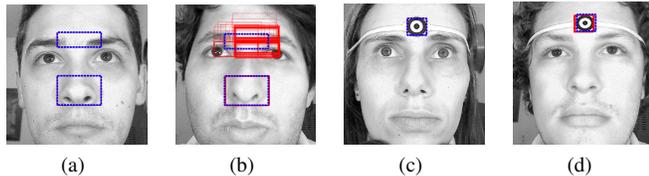


Fig. 14. Exemplos de resultados da detecção e rastreamento do ponto fixo para (a-b) nariz, sobrancelha e (c-d) marcação artificial (melhor e pior caso). Os retângulos tracejados indicam a detecção do ponto fixo no primeiro quadro de vídeo e os demais indicam o seu rastreamento nos quadros seguintes

Nota-se uma boa estabilidade no rastreamento da marcação artificial, pois é um ponto fixo bastante distinto e de fácil localização. O nariz apresentou uma boa estabilidade, porém sua localização varia em relação aos olhos quando o usuário movimentava a cabeça, o que prejudica a estimação da direção do olhar. As sobrancelhas tiveram uma instabilidade alta no rastreamento pois são elementos inconstantes na imagem devido principalmente à sua movimentação pelos usuários. Portanto, não é recomendável o seu uso como ponto fixo.

C. Estimação da direção do olhar

Para cada ponto estimado na fase de teste, foi calculada sua distância euclidiana do ponto real e foi tirada a média dessas distâncias. Esse valor indica o erro da estimação e foi calculado para cada usuário. A medida de acurácia mais

utilizada na literatura é o acerto do sistema em graus [2], neste trabalho denominado de cone de erro. A Fig. 15 ilustra o cone de erro. O raio do cone equivale ao erro médio da estimação.

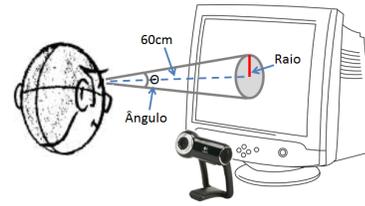


Fig. 15. Ilustração do cone de erro

O cálculo do ângulo é feito através da equação 1.

$$\Theta = 2 \times \arctan \left(\frac{\text{erro da estimação}}{60\text{cm}} \right) \quad (1)$$

A tabela II exhibe os resultados da estimação do olhar em graus para cada usuário da base de dados e cada um dos três pontos fixos. Quanto menor o valor, melhor é a estimação.

TABELA II

RESULTADOS DA ESTIMAÇÃO DO OLHAR UTILIZANDO CADA UM DOS TRÊS PONTOS FIXOS. A COLUNA "MELHOR RESULTADO" INDICA A SITUAÇÃO ONDE OCORREU A MELHOR PRECISÃO ATRAVÉS DO PAR X/Y, ONDE X PODE SER N=NARIZ, MA=MARCAÇÃO ARTIFICIAL E S=SOBRANCELHAS, E Y PODE SER S=SEM O REFINAMENTO E C=COM O REFINAMENTO.

Usuário	Sem o refinamento			Com o refinamento			Melhor resultado
	Nariz	Sobrancelhas	Marc. artificial	Nariz	Sobrancelhas	Marc. artificial	
1	7.33°	6.30°	7.54°	6.69°	5.63°	6.19°	S/C
2	9.70°	8.92°	6.80°	8.36°	6.88°	4.64°	MA/C
3	14.06°	13.36°	7.55°	13.20°	11.79°	4.84°	MA/C
4	5.53°	30.57°	5.48°	3.62°	30.20°	6.37°	N/C
5	15.25°	37.66°	7.85°	15.04°	37.31°	5.90°	MA/C
Média	10.37°	19.36°	7.05°	9.38°	18.36°	5.59°	MA/C

Na grande maioria dos casos os melhores resultados foram obtidos com o refinamento da localização da íris. Quanto ao ponto fixo, a marcação artificial gerou os melhores resultados. Os piores resultados foram obtidos nos usuários 4 e 5, usando a sobrancelha como ponto fixo. Isso se deve ao fato desses usuários não possuírem sobrancelhas muito acentuadas, o que indica que o uso destas como ponto fixo é desaconselhável.

O uso do nariz como ponto fixo apresentou bons resultados, principalmente para o usuário 4, obtendo o melhor resultado. Porém, este ponto fixo gerou os piores resultados para o usuário 3. Isso se deve pelo fato do mesmo ter movimentado o nariz durante a captura das imagens. A marcação artificial parece ser o ponto fixo mais estável para ser utilizado.

Na Fig. 16 pode-se visualizar graficamente os resultados obtidos. Cada retângulo representa uma captura, contendo os 25 pontos de teste. Para cada ponto, estão representados os pontos referentes à estimativa do sistema nos cinco quadros de vídeo capturados para cada usuário. Em cada retângulo, estão agrupados os dados dos cinco usuários. Nota-se um

melhor agrupamento dos pontos nas capturas realizadas com o refinamento, como foi argumentado na seção V-A.

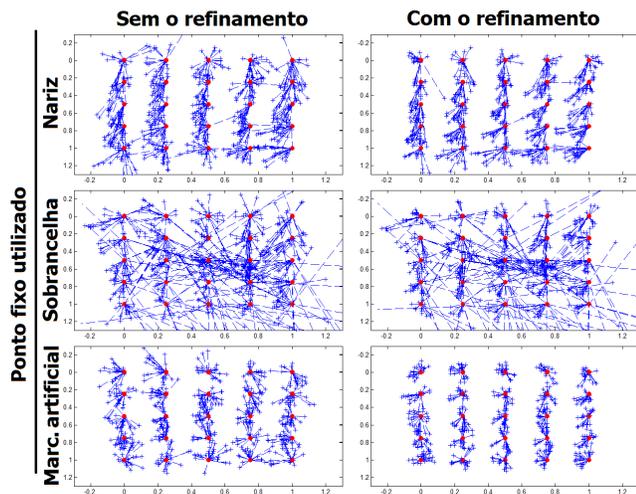


Fig. 16. Visualização gráfica dos resultados da estimativa da direção olhar

VI. CONCLUSÕES

Neste trabalho foi feito um estudo sobre rastreamento do olhar e desenvolvido o protótipo de um sistema de detecção da direção do olhar. O protótipo foi desenvolvido de forma modular, para melhor dividir as etapas típicas existentes nesse tipo de sistema.

A principal contribuição deste trabalho foi ter conseguido resultados significativos utilizando uma *webcam* comum ao invés de uma câmera mais robusta, e usando iluminação natural ao invés de infravermelha, o que diminui os custos.

REFERÊNCIAS

- [1] R. G. Monteiro, “Detecção da direção do olhar via *webcam*,” Dissertação de Mestrado, PPGI/UFRJ, Rio de Janeiro, RJ, Brasil, 2012.
- [2] D. HANSEN and Q. JI, “In the eye of the beholder: A survey of models for eyes and gaze,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 478–500, Março 2010.
- [3] Y. TIAN, T. KANADE, and J. F. COHN, “Dual-state parametric eye tracking,” *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition*, 2000.
- [4] D. HANSEN and A. E. C. PECE, “Eye tracking in the wild,” *Computer Vision and Image Understanding*, vol. 98, no. 1, pp. 182–210, Abril 2005.
- [5] J. HUANG and H. WECHSLER, “Eye detection using optimal wavelet packets and radial basis functions (rbfs),” *International Journal of Pattern recognition and Artificial Intelligence*, vol. 13, no. 7, 1999.
- [6] W. HUANG and R. MARIANI, “Face detection and precise eyes location,” *Proceedings of the International Conference on Pattern Recognition (ICPR’00)*, 2000.
- [7] D. HANSEN, J. P. HANSEN, M. NIELSEN, A. S. JOHANSEN, and M. B. STEGMANN, “Eye typing using markov and active appearance models,” *IEEE Workshop on Applications on Computer Vision*, pp. 132–136, 2003.
- [8] T. ISHIKAWA, S. BAKER, I. MATTHEWS, and T. KANADE, “Passive driver gaze tracking with active appearance models,” *Proceedings of the 11th World Congress on Intelligent Transportation Systems*, Outubro 2004.
- [9] Y. MATSUMOTO and A. ZELINSKY, “An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement,” *International Conference on Automatic Face and Gesture Recognition*, pp. 499–504, 2000.
- [10] X. L. C. BROLLY and J. B. MULLIGAN, “Implicit calibration of a remote gaze tracker,” *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW’04)*, vol. 8, p. 134, 2004.
- [11] Y. EBISAWA and S. SATOH, “Effectiveness of pupil area detection technique using two light sources and image difference method,” *Proceedings of the 15th Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society, San Diego, CA*, pp. 1268–1269, 1993.
- [12] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, 1999.
- [13] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2001, pp. 511–518.
- [14] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *ICCV*, 1998, pp. 839–846.
- [15] I. Sobel and G. Feldman, “A 3x3 isotropic gradient operator for image processing,” 1968, presented at a talk at the Stanford Artificial Project, never published.
- [16] J. Canny, “A computational approach to edge detection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 6, pp. 679–698, November.
- [17] P. V. C. Hough, “Machinex analysis of bubble chamber pictures,” in *Proceedings of International Conference on High Energy Accelerators and Instrumentation*, 1959.
- [18] R. O. Duda and P. E. Hart, “Use of the hough transformation to detect lines and curves in pictures,” *Commun. ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [19] B. Kunka and B. Kostek, “Non-intrusive infrared-free eye tracking method,” in *Signal Processing Algorithms, Architectures, Arrangements, and Applications Conference Proceedings (SPA)*, 2009, sept. 2009, pp. 105–109.
- [20] B. Zitova and J. Flusser, “Image registration methods: a survey,” *Imag. Vis. Comput.*, vol. 21, pp. 977–1000, 2003.
- [21] O. Jesorsky, K. J. Kirchberg, and R. Frischholz, “Robust face detection using the hausdorff distance,” in *Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication*, ser. AVBPA ’01, 2001, pp. 90–95.