

Identificação de Eventos do Olhar: Métodos

Universidade de São Paulo

Introdução

Rastreamento do olhar é uma área que estuda o comportamento do olhar, analisando o processamento visual e cognitivo da pessoa. Os movimentos do olhar são coletados por equipamentos denominados rastreadores de olhar (*eye trackers*) para análise, comparação e para desenvolver sistemas mais interativos.

Dependendo da taxa de amostragem do equipamento e da aplicação pretendida, a quantidade de dados coletados poder ser muito grande para sua análise. Nesse caso é necessária uma transformação destes dados em eventos que informem melhor o processamento visual e cognitivo durante o experimento.

Os rastreadores basicamente informam a coordenada bidimensional de cada ponto coletado e seu *timestamp*. Outras propriedades também podem ser informadas, como o diâmetro da pupila. As amostras configuram uma série temporal. Os dados de tempo e espaço podem ser usados para analisar os eventos.

Eventos

Os eventos do movimento do olhar podem ser:

- **Fixações:** Para Salvucci [1] e Karrsgard[2], uma fixação é uma pausa sobre regiões de interesse.
- **Sacadas:** Um rápido movimento entre fixações [1];
- **Movimentos fixacionais:** O olho não fica parado durante uma fixação, todavia ocorrem pequenos movimentos, como *drifts* (desvios do foco para fora do alvo durante a fixação), tremores (rápidos, de baixa amplitude e involuntários), microssacadas (correções dos *drifts* e renovação do estímulo na retina) e nystagmus (patologia periódica composta por movimentos suaves e rápidos alternadamente, que pode provocar tontura e sensação de movimento em objetos estáticos).
- **Perseguição Contínua:** O movimento do olho acompanhando um alvo em movimento é denominado *perseguição contínua*, *perseguição suave*, ou simplesmente *perseguição*.

- **Glissadas:** No fim da sacada, geralmente o olho não para no ponto desejado, e sim ele o passa. Dessa forma, ele faz um pequeno movimento senoidal para corrigir e fixar-se no ponto. Esse movimento é denominado *glissada*.

Filtros

A análise de dados do olhar é dividida em duas partes: filtragem e classificação. Os ruídos também são divididos em duas categorias:

- Ruídos provenientes do equipamento;
- Movimentos do olhar que não estão sendo analisados.

Os filtros servem para remover, ou pelo menos reduzir, o primeiro tipo de ruído. A classificação elimina o segundo tipo. Essa separação em duas etapas faz com que os métodos de classificação sejam mais independentes dos equipamentos e suas características.

Alguns filtros utilizados na literatura são:

- Filtro de Resposta ao Impulso Finita (FIR) [3];
- Filtro de Média [4];
- Filtro de Mediana [4];
- Filtro Savitzky-Golay [5];
- Filtro de Kalman [6].

Categorias de Métodos de Classificação

Salvucci [1] introduz uma taxonomia de algoritmos de identificação de fixações e sacadas. Esta taxonomia é baseada em como são usadas as informações de tempo e espaço. Os algoritmos citados no artigo representam classes de técnicas que compartilham algum critério de identificação. Ele também apresenta uma forma de analisar os algoritmos de maneira qualitativa: facilidade de uso, velocidade de interpretação, acurácia, robustez, e parametrização.

Critérios espaciais

- Baseados em velocidade: Estes algoritmos utilizam o fato dos pontos que compõem uma fixação terem uma velocidade baixa, enquanto que numa sacada, as velocidades dos pontos são altas. Geralmente são utilizadas para classificar sacadas;

- Baseados em dispersão: Em uma fixação, os pontos são próximos entre si. Medidas de dispersão podem ser utilizadas para classificar fixações;
- Baseados em área de interesse (AOI): Os pontos são agrupados de acordo com regiões pré-determinadas pela aplicação. Servem geralmente para uma análise de alto nível. Podem ser usados algoritmos baseados em velocidade e/ou dispersão antes de realizar a análise baseado em AOI.

Crítérios temporais

- Sensível à duração: Utiliza uma duração mínima para descartar fixações com uma duração curta demais para os limites fisiológicos do olho.
- Localmente adaptativo: pontos vizinhos influenciam a classificação de um ponto específico. Robusto contra ruídos.

Métodos

Os algoritmos propostos por Salvucci representam as características mais básicas e distintas dos algoritmos criados e publicados antes do seu artigo de revisão: I-VT, I-DT, I-HMM, I-MST, I-AOI.

I-VT

Este algoritmo representativo proposto por Salvucci é um dos mais básicos. Contém um parâmetro, o limiar de velocidade. Recebendo as amostras dos pontos, calculam-se suas velocidades. Se a velocidade for menor que o limiar, o respectivo ponto é classificado como fixação, senão é classificado como sacada.

Vantagens

- fácil de implementar;
- eficiente;
- pode ser executado em tempo real.

Desvantagens

- instável em pontos com velocidade próxima do threshold (precisa lidar com o ruído do equipamento e movimentos do olhar irrelevantes para a pesquisa);
- Pode provocar alternâncias entre classificações, implicando em fixações e sacadas com poucos pontos, aumentando o número de fixações excluídas pelo critério de duração mínima;

- Não é robusto;
- Perseguições podem ser classificados como fixações ou sacadas dependendo de sua velocidade.

I-HMM

Este algoritmo utiliza uma máquina de 2 estados para classificar fixação e sacada, recebendo parâmetros de distribuição das velocidades (média e desvio padrão para cada estado), além das probabilidades de transição entre estados. O modelo pode ser treinado para reestimar os parâmetros.

Vantagens:

- Modelo probabilístico ao invés de um threshold. Utiliza informação sequencial (os vizinhos influenciam o ponto);
- É mais robusto na presença do ruído;
- Pode expandir o diagrama de estados (incorporando mais movimentos do olhar);
- É executado em tempo linear e pode ser executado em tempo real.

Desvantagens:

- Mais complexo que I-VT;
- Procedimento de reestimar os parâmetros também é complexo.

I-DT

Este algoritmo utiliza o critério de dispersão para agrupar os pontos em uma fixação. Ele inicia uma janela com tamanho de acordo com a duração mínima de uma fixação (fixações curtas são descartadas), geralmente 100 ms. Caso a medida de dispersão dos pontos dentro da janela for menor que um limiar, então a janela é expandida até que a dispersão seja maior, agrupando todos os pontos na janela como uma fixação. Salvucci utilizou como critério de dispersão $(Max_x - Min_x) + (Max_y - Min_y)$. Outras medidas de dispersão podem ser usadas:

- Distância entre qualquer um dos pontos;
- Distância entre os pontos e o centro da fixação;
- Desvio padrão das coordenadas.

Vantagens:

- Algoritmo simples
- Tempo linear;
- Pode ser feito em tempo real;
- Resultado parecido com a saída do I-HMM (sendo mais robusto do que o I-VT).

Desvantagens:

- Parâmetros interdependentes (ex: duração mínima alta e limiar de dispersão baixa pode não classificar nenhuma fixação);
- Sensível a ruído no critério espacial (pode ultrapassar o limiar);
- Possíveis fixações dispersas podem não ser classificadas.

I-MST

Este algoritmo cria uma estrutura de árvore que interliga os pontos de tal forma que a soma dos comprimentos das arestas da árvore seja o menor possível. Para construir a árvore, utiliza-se o algoritmo de Prim . É localmente adaptativo por interligar os pontos aos seus vizinhos, direta ou indiretamente.

Vantagens:

- Robusto (pode usar variância e média para lidar com ruído);
- Cria clusters de fixações
- Podem-se usar outras caracterizações que não sejam meramente espaciais para classificar as fixações.

Desvantagens:

- Lento (tempo de execução exponencial);
- Para cada ponto adicionado, é necessário achar o ponto mais próximo dentre vários para reestruturar o cluster e separar os clusters.

I-AOI

Este método de classificação de alto nível converte as amostras em regiões de acordo com divisões da região do estímulo. Cada região é identificada com um símbolo. O resultado do método transforma as amostras em uma sequência de símbolos, cujas sequências podem ser comparadas entre si usando o algoritmo de Levenshtein[7].

Vantagens:

- Tempo real;
- Simples;
- Análise de alto nível.

Desvantagens:

- Não lida bem com sacadas (incluídas nas fixações se estiverem dentro das regiões), aumentando a duração da fixação;
- Longas sacadas são consideradas fixações nas regiões intermediárias;
- Depende da aplicação (distribuição das regiões).

Avaliação dos Métodos

Shic[8] explora diferentes algoritmos de identificação de fixações mostrando que suas interpretações podem ser diferentes, mesmo trabalhando com os mesmos dados coletados. Ele analisa os seguintes algoritmos baseados em dispersão:

- Dispersão de Distância: a distância entre dois pontos quaisquer na fixação não pode superar um limiar. É executado em $O(n^2)$;
- Centróide: os pontos de uma fixação não podem ser mais distantes do que um limiar para sua centróide. Pode construir uma versão em tempo real, computando apenas os novos pontos;
- Posição-Variância: modela o grupo de pontos como uma distribuição gaussiana, e não podem ultrapassar um desvio padrão de distância;
- I-DT de Salvucci: a soma da máxima distância horizontal com a máxima distância vertical deve ser menor que um limiar.

Ele viu que o tempo de fixação médio segue um comportamento linear para valores que correspondem aos limites fisiológicos da visão foveal (desvio padrão da dispersão até 1° e tempo mínimo de fixação até 200ms), mesmo que o número de fixações e o total de tempo gasto nas fixações forem não lineares.

Salvucci [1] avalia os algoritmos de acordo com critérios subjetivos:

- Facilidade de implementação;
- Acurácia;
- Velocidade;
- Robustez;
- Número de parâmetros.

O único critério quantitativo é o número de parâmetros, visto que ele definiu os outros critérios qualitativamente, embora possam ser criadas métricas para torná-los objetivos. Também não há um valor final devido à subjetividade, todavia também pode ser criado um cálculo usando e agregando os critérios de forma quantitativa.

Larsson [9] em sua tese apresenta um método de avaliação denominado *Precision and Recall*. O método classifica a saída dos algoritmos em 4 tipos baseados na *predição* (a saída do algoritmo) e no *padrão de ouro* (a correta classificação):

- Padrão de ouro como *Sim* e Predição como *Sim*: Verdadeiro Positivo (VP);
- Padrão de ouro como *Sim* e Predição como *Não*: Falso Positivo (FP);
- Padrão de ouro como *Não* e Predição como *Sim*: Falso Negativo (FN);
- Padrão de ouro como *Não* e Predição como *Não*: Verdadeiro Negativo (VN).

O objetivo da etapa *Precision* é saber a razão entre os verdadeiros positivos – o algoritmo classificou corretamente como *Sim* – e todos os classificados como *Sim* pelo algoritmo, mesmo os falso positivos. O objetivo da etapa *Recall* é saber a razão entre os verdadeiros positivos e todos que deveriam ser classificados como *Sim*, de acordo com o padrão de ouro.

Precision =

$$\frac{VP}{VP + FP}$$

Recall =

$$\frac{VP}{VP + FN}$$

Conclusão

Esta revisão serve para conhecer os métodos de classificação de dados do olhar e métodos de avaliação dos algoritmos de análise. Como trabalho futuro, as etapas de filtragem, bem como outros métodos de classificação, serão descritas e analisadas. Classificações de outros eventos do olhar também serão levadas em conta, como piscadas, perseguições contínuas e nystagmus. Alguns dos métodos que estão sendo avaliados são:

- HMM de Karrsgard [2];
- Clusterização de Projeção de Urruty [10];
- Variância e Covariância de Veneri [11];
- *Mean Shift Procedure* de Santella [12].

Bibliografia

- [1] D. D. Salvucci and J. H. Goldberg, “Identifying fixations and saccades in eye-tracking protocols,” in *Proceedings of the symposium on Eye tracking research & applications - ETRA '00*, 2000, pp. 71–78.
- [2] I. Kärregård and A. Lindholm, “Eye movement tracking using hidden Markov models,” PhD thesis.
- [3] P. Olsson, “Real-time and Offline Filters for Eye Tracking,” PhD thesis, KTH Electrical Engineering; KTH Electrical Engineering, Stockholm, Sweden, 2007.
- [4] A. Olsen, “The Tobii I-VT Fixation Filter,” 2012.
- [5] M. Nyström and K. Holmqvist, “An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data,” *Behavior research methods*, vol. 42, no. 1, pp. 188–204, Feb. 2010.
- [6] D. Sauter, B. J. Martin, N. Renzo, and C. Vomscheid, “Analysis of eye tracking movements using innovations generated by a Kalman filter,” *Medical & Biological Engineering & Computing*, vol. 29, no. 1, pp. 63–69, Jan. 1991.
- [7] F. Cristino, S. Mathôt, J. Theeuwes, and I. D. Gilchrist, “ScanMatch: a novel method for comparing fixation sequences,” *Behavior research methods*, vol. 42, no. 3, pp. 692–700, Aug. 2010.
- [8] F. Shic, B. Scassellati, and K. Chawarska, “The incomplete fixation measure,” in *Proceedings of the 2008 symposium on Eye tracking research & applications - ETRA '08*, 2008, p. 111.
- [9] G. Larsson, “Evaluation methodology of eye movement classification algorithms,” PhD thesis, Royal Institute of Technology, 2010.
- [10] T. Urruty, S. Lew, N. Ihadaddene, and D. A. Simovici, “Detecting eye fixations by projection clustering,” *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 3, no. 4, pp. 1–20, Dec. 2007.
- [11] G. Veneri, P. Piu, F. Rosini, P. Federighi, A. Federico, and A. Rufa, “Automatic eye fixations identification based on analysis of variance and covariance,” *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1588–1593, Oct. 2011.
- [12] A. Santella and D. DeCarlo, “Robust clustering of eye movement recordings for quantification of visual interest,” in *Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications - ETRA '2004*, 2004, pp. 27–34.