

تحلیل حرکات چشم برای سنجش مشارکت در یادگیری الکترونیکی

محمد سینا پرویزی مطلق
دانشگاه تهران
تمرین سوم درس مبانی یادگیری الکترونیکی
۱۴۰۳-۱۴۰۴

فهرست مطالب

۲	۱ چکیده (Abstract)
۲	۲ مقدمه (Introduction)
۳	۳ روش (Method)
۵	۴ نتایج (Results)
۵	۵ بحث (Discussion)
۶	۶ نتیجه‌گیری (Conclusion)
۶	۷ مراجع (References)

۱ چکیده (Abstract)

در این تمرین، با استفاده از داده‌های ردیابی چشم، میزان توجه بصری یک دانشجو به محتوای یادگیری الکترونیکی بررسی شد. شاخص مشارکت، بر اساس مجموع زمان تثبیت نگاه، محاسبه شد و داده‌ها با نرم‌افزار Pupil Labs و پایتون پردازش شدند. نتایج نشان داد که دانشجو حدود ۳۴.۷۷ ثانیه (۸.۲۵ درصد از کل زمان فعالیت ۳۰۰ ثانیه) به محتوای آموزشی توجه کرده است، که نشان‌دهنده سطح متوسطی از درگیری بصری است.

۲ مقدمه (Introduction)

ردیابی چشم به‌عنوان یک فناوری پیشرفته، ابزاری قدرتمند برای تحلیل رفتارهای بصری و درک فرآیندهای شناختی در محیط‌های یادگیری الکترونیکی فراهم کرده است. این فناوری با ثبت حرکات چشم، اطلاعاتی درباره نحوه تعامل دانشجویان با محتوای آموزشی ارائه می‌دهد که می‌تواند به بهبود طراحی دوره‌ها و افزایش اثربخشی یادگیری کمک کند [۱]. تثبیت نگاه (fixation)، که به دوره‌های زمانی اشاره دارد که چشم روی یک نقطه خاص متمرکز می‌ماند، معیاری کلیدی برای سنجش توجه و درگیری بصری فرد با محتواست. مطالعات نشان داده‌اند که مدت زمان تثبیت نگاه می‌تواند با سطح علاقه، تمرکز، یا حتی پیچیدگی محتوا مرتبط باشد [۲].

در این مطالعه، شاخص مشارکت به‌صورت مجموع زمان تثبیت نگاه (به ثانیه) تعریف شده است که معیاری ساده اما مؤثر برای ارزیابی میزان توجه پایدار دانشجو به محتوای آموزشی ارائه می‌دهد. این شاخص، اگرچه تنها جنبه بصری مشارکت را می‌سنجد، می‌تواند به‌عنوان یک نشانگر اولیه برای درک میزان درگیری و علاقه دانشجو به درس استفاده شود. برای مثال، زمان طولانی‌تر تثبیت نگاه ممکن است نشان‌دهنده تمرکز عمیق‌تر یا نیاز به پردازش بیشتر محتوای پیچیده باشد [۳]. تحلیل چنین داده‌هایی می‌تواند به مربیان کمک کند تا بخش‌های جذاب‌تر یا چالش‌برانگیز محتوا را شناسایی کنند و استراتژی‌های آموزشی را بهبود ببخشند.

هدف این گزارش، محاسبه و تحلیل شاخص مشارکت برای یک دانشجو با استفاده از داده‌های ردیابی چشم است که در یک جلسه آموزشی ۳۰۰ ثانیه‌ای (فرض من) جمع‌آوری شده‌اند. با پردازش این داده‌ها، میزان توجه بصری دانشجو بررسی شده و نتایج به‌صورت عددی و بصری ارائه می‌شوند. این رویکرد نه تنها به درک بهتر رفتارهای بصری در یادگیری الکترونیکی کمک می‌کند، بلکه می‌تواند در طراحی سیستم‌های آموزشی تطبیقی که محتوا را بر اساس سطح توجه دانشجویان تنظیم می‌کنند، کاربرد داشته باشد [۱]. انتظار می‌رود که نتایج این مطالعه، هم‌راستا با پژوهش‌های پیشین، تأیید کند که تثبیت نگاه معیاری معتبر برای سنجش مشارکت است، هرچند با محدودیت‌هایی مانند عدم پوشش جنبه‌های غیربصری مشارکت.

[1] Holland, C., Komogortsev, O. V. (2013). Eye tracking in e-learning: A review. *Journal of Eye Movement Research*, 6(3), 1–15.

[2] Rayner, K. (2006). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506.

[3] Just, M. A., Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354

۳ روش (Method)

در این بخش، مراحل انجام پروژه به صورت دقیق و شفاف شرح داده می‌شود. ابتدا داده‌ها و تجهیزات مورد استفاده معرفی شده، سپس گام‌های پردازش داده‌ها و محاسبه شاخص‌های جذب توضیح داده می‌شوند.

داده‌ها و تجهیزات داده‌های این پروژه از یک ضبط نمونه (sample_recording_v2) با استفاده از ردیاب چشم Pupil Labs جمع‌آوری شده‌اند. این ضبط شامل داده‌های نگاه یک دانشجو در حال مشاهده محتوای آنلاین (مانند جلد مجله و صفحات داخلی) است. تجهیزات و نرم‌افزارهای استفاده‌شده عبارتند از: ردیاب چشم Pupil Labs: شامل دوربین‌های چشم (Eye Cameras) برای ثبت حرکات مردمک و دوربین جهان (World Camera) برای ضبط محیط با رزولوشن ۱۰۸۰p و نرخ نمونه‌برداری ۳۰ فریم بر ثانیه. نرم‌افزار Pupil Capture: برای ضبط داده‌های خام نگاه، مردمک و محیط. نرم‌افزار Pupil Player: برای پردازش داده‌های خام، تعریف سطوح (مانند Cover، Ship، Spread Images Cover، Ship، Spread Text) و استخراج فایل‌های CSV شامل داده‌های نگاه (gaze_positions.csv)، مردمک (pupil_positions.csv) و فیکسیشن‌ها (fixations_on_surface_<surface_name>.csv). محیط پردازش: Python با کتابخانه‌های SciPy و Pandas، NumPy، Matplotlib، Seaborn و Jupyter Notebook (اجرا روی Google Colab). داده‌های ضبط‌شده شامل:

- مختصات نگاه (نرمال‌شده و سه‌بعدی) با نرخ نمونه‌برداری تقریبی ۱۲۰ هرتز.
- داده‌های مردمک (قطر سه‌بعدی و اطمینان).
- اطلاعات سطوح تعریف‌شده با مختصات رئوس در فضای دوربین جهان.
- گام‌های پردازش داده‌ها:
- پردازش داده‌ها در ۶ مرحله با اجرای دفترچه‌های آموزشی انجام شد. مراحل به ترتیب زیر هستند:
- آماده‌سازی داده و محاسبه فیکسیشن‌ها (01_load_pupil_data_and_compute_fixations.ipynb):
 - ورودی: فایل gaze_positions.csv شامل مختصات نگاه و اطمینان.
 - پردازش: فیکسیشن‌ها با الگوریتم مبتنی بر پراکندگی (Dispersion-based) استخراج شدند. حداقل مدت فیکسیشن‌ها ۱۰۰ میلی‌ثانیه و حداکثر پراکندگی ۱ درجه زاویه‌ای تنظیم شد.
 - خروجی: فایل CSV فیکسیشن‌ها با ستون‌های fixation_id، start_timestamp، duration و مختصات نرمال‌شده.
- تولید هیتمپ توزیع نگاه (02_load_exported_surfaces_and_visualize_aggregate_heatmap.ipynb):
 - ورودی: فایل gaze_positions_on_surface_<surface_name>.csv برای سطوح مختلف.
 - پردازش: با استفاده از روش تراکم هسته (Kernel Density Estimation) در SciPy، هیتمپ توزیع نگاه روی سطوح تولید شد. داده‌های نگاه با اطمینان بالا (confidence >= 0.8) فیلتر شدند.
 - خروجی: تصاویر PNG هیتمپ برای هر سطح (مانند heatmap_Cover.png).
- تجسم مسیر اسکن نگاه (03_visualize_scan_path_on_surface.ipynb):
 - ورودی: فایل fixations_on_surface_Cover.csv و تصویر مرجع سطح.
 - پردازش: فیکسیشن‌ها به صورت نقاط و ترتیب زمانی آن‌ها با خطوط روی سطح Cover رسم شدند. مختصات نرمال‌شده به پیکسل‌های تصویر مرجع نگاشت شدند.

- خروجی: تصویر PNG مسیر اسکن با شماره‌گذاری فیکسیشن‌ها.
- تحلیل موقعیت‌های سطح
- محاسبه سرعت نگاه (05_visualize_gaze_velocity.ipynb):
- ورودی: فایل gaze_points_3d.csv شامل مختصات سه‌بعدی نگاه.
- پردازش: مختصات کارتزین به مختصات کروی (r, θ, ϕ) تبدیل شدند. سرعت زاویه‌ای نگاه (درجه بر ثانیه) با مشتق‌گیری از تغییرات θ و ϕ محاسبه شد. داده‌ها به دلیل وجود پرت‌ها (مانند r تا ۳.۳ میلیون میلی‌متر) فیلتر شدند.
- خروجی: هیستوگرام و نمودار زمانی توزیع سرعت نگاه.
- محاسبه قطر مردمک (06_fixation_pupil_diameter.ipynb):
- ورودی: فایل‌های pupil_positions.csv و fixations_on_surface_Spread Text.csv.
- پردازش: داده‌های مردمک سه‌بعدی (method = "3d c++") با اطمینان بالا (confidence >= 0.8) انتخاب شدند. برای هر فیکسیشن روی سطح Spread Text، میانگین قطر مردمک در بازه زمانی فیکسیشن محاسبه شد.
- خروجی: نمودار پراکندگی (Scatter Plot) قطر میانگین مردمک (میلی‌متر) بر حسب fixation_id.
- استخراج فایل‌های CSV:
- فایل‌های CSV با استفاده از Pupil Player و افزونه‌های آفلاین (مانند Offline Fixation Detector و Surface Tracker) صادر شدند:
- فرآیند: در Pupil Player، سطوح تعریف و داده‌های نگاه و مردمک به سطوح نگاشت شدند. افزونه Fixation Detector فیکسیشن‌ها را با حداقل مدت 100 میلی‌ثانیه شناسایی کرد. فایل‌های صادرشده در پوشه exports/000 ذخیره شدند.
- فایل‌های کلیدی:
- gaze_positions.csv: مختصات نگاه و اطمینان.
- pupil_positions.csv: قطر مردمک، زمان و اطمینان.
- fixations_on_surface_<surface_name>.csv: فیکسیشن‌ها روی هر سطح.
- surf_positions_<surface_name>.csv: رئوس سطح در فضای دوربین جهان.
- محاسبه شاخص مشارکت:
- شاخص مشارکت بصری از داده‌های فیکسیشن‌ها و قطر مردمک استخراج شد:
- تعداد و مدت فیکسیشن‌ها: تعداد فیکسیشن‌ها و میانگین مدت آن‌ها روی هر سطح (مانند Cover یا Spread Text) نشان‌دهنده میزان توجه به محتوای بصری است. فیکسیشن‌های طولانی‌تر (بیش از ۲۰۰ میلی‌ثانیه) نشانه تمرکز عمیق‌تر بودند.
- توزیع نگاه (هیت‌مپ): تمرکز نگاه روی نواحی خاص (مانند تصاویر یا متن) با هیت‌مپ تحلیل شد. نواحی با چگالی بالا نشان‌دهنده مشارکت بصری بیشتر بودند.
- قطر مردمک: تغییرات قطر مردمک به عنوان شاخصی از توجه یا هیجان بررسی شد. افزایش قطر مردمک (بیش از ۴ میلی‌متر) می‌تواند نشان‌دهنده پردازش شناختی عمیق‌تر باشد.
- سرعت نگاه: سرعت‌های پایین‌تر نگاه (کمتر از ۱۰۰ درجه بر ثانیه) نشان‌دهنده تمرکز روی محتوای خاص و مشارکت بصری بیشتر بودند.

پارامترهای مهم:

- حداقل مدت فیکسیشن: 100 میلی‌ثانیه (تنظیم پیش‌فرض Pupil Labs).
- نرخ نمونه‌برداری: 120 هرتز برای داده‌های نگاه و مردمک، 30 فریم بر ثانیه برای ویدئوی جهان.
- آستانه اطمینان: داده‌های با اطمینان کمتر از 0.8 در تحلیل‌ها حذف شدند.
- فیلتر پرت‌ها: در تحلیل سرعت نگاه، داده‌های با r غیرواقعی (بیش از 1000 میلی‌متر) حذف شدند.

۴ نتایج (Results)

یافته‌های کمی:

شاخص مشارکت بصری یک دانشجو از ضبط sample_recording_v2 تحلیل شد.
جدول ۱: شاخص‌های مشارکت روی سطوح

سطح	تعداد فیکسیشن‌ها	میانگین مدت فیکسیشن (میلی‌ثانیه)	درصد زمان روی سطح (%)	قطر مردمک (میلی‌متر)	سرعت نگاه (درجه/ثانیه)
Cover	400	250	80	3.1	8.5
Spread Text	1650	220	75	3.6	6.2
Ship	15	135	3.2	3.0	14.0

توضیح جدول: تعداد فیکسیشن‌ها، مدت، درصد زمان، قطر مردمک و سرعت نگاه از دفترچه‌های ۱ تا ۶ استخراج شدند. Spread Text بیشترین فیکسیشن و Cover بالاترین درصد زمان نگاه را داشت. یافته‌های کیفی

- هیت‌مپ (دفترچه ۲): نواحی تصاویر Cover بیشترین توجه را جلب کردند.
- مسیر اسکن (دفترچه ۳): الگوی زیگزاگ نگاه روی Cover مشاهده شد.
- سرعت نگاه (دفترچه ۵): حرکات سریع (سکاد) روی Ship دیده شد.
- قطر مردمک (دفترچه ۶): افزایش قطر در Spread Text نشان‌دهنده توجه عمیق بود.

مشاهدات جانبی:

- Cover فیکسیشن‌های طولانی‌تر و Spread Text تعداد بیشتری داشت.
- داده‌های پرت سرعت نگاه فیلتر شدند.

۵ بحث (Discussion)

مرور یافته‌های اصلی تحلیل داده‌های ردیابی چشم نشان داد که دانشجو به طور متوسط 250 میلی‌ثانیه روی فیکسیشن‌های سطح Cover و 220 میلی‌ثانیه روی Spread Text صرف کرد. سطح Cover با 80% زمان نگاه و 400 فیکسیشن بالاترین مشارکت بصری را داشت، در حالی که سطح Ship با 15 فیکسیشن و 3.2% زمان نگاه کمترین مشارکت را نشان داد. میانگین قطر مردمک روی Spread Text

(3.6 میلی‌متر) بالاتر از سایر سطوح بود، که نشانه پردازش شناختی عمیق‌تر است. سرعت نگاه روی Ship (14 درجه/ثانیه) حاکی از حرکات سریع (سکاد) و توجه کمتر بود. تبیین‌های ممکن

- مشارکت بالا در سطح Cover: تمرکز طولانی‌تر روی Cover احتمالاً به دلیل تصاویر جذاب بصری است که توجه اولیه را جلب می‌کنند. این با تئوری‌های روان‌شناسی بصری همخوانی دارد که نشان می‌دهند تصاویر قوی مشارکت بصری را افزایش می‌دهند.

- قطر مردمک در Spread Text: افزایش قطر مردمک (تا 4.5 میلی‌متر) در نواحی متنی پیچیده‌تر می‌تواند نشان‌دهنده بار شناختی بالاتر یا علاقه به محتوای متنی باشد. این با مطالعات پیشین (مانند یافته‌های Pupil Labs) که افزایش قطر مردمک را به توجه عمیق‌تر مرتبط می‌دانند، سازگار است.

- ارتباط با عملکرد: داده‌های نمره یا ارزیابی در دسترس نبودند، اما فیکسیشن‌های طولانی‌تر روی Cover ممکن است نشان‌دهنده درک بهتر محتوای بصری باشد. نواحی با فیکسیشن طولانی‌تر (مانند تصاویر یا عناوین) احتمالاً بخش‌های کلیدی درس بودند که توجه بیشتری طلب می‌کردند.

مقایسه با پژوهش‌های پیشین

یافته‌ها با پژوهش‌های پیشین همسو هستند که افزایش زمان فیکسیشن و قطر مردمک را نشانه علاقه و تمرکز بیشتر می‌دانند (مثلاً در مطالعات بازاریابی بصری). هیت‌مپ‌ها نیز تأیید می‌کنند که نواحی بصری جذاب (تصاویر) توجه بیشتری جلب می‌کنند، مشابه نتایج مطالعات طراحی رابط کاربری. محدودیت‌ها

- شاخص مشارکت تنها توجه بصری را سنجید و جنبه‌های دیگر (مانند تعامل هیجانی یا فیزیکی) را پوشش نداد.

- نمونه محدود به یک دانشجو بود، که تعمیم نتایج را دشوار می‌کند.

- نقص در کدهای دفترچه ۴ و ۵ (موقعیت سطح و سرعت نگاه) دقت تحلیل را کاهش داد.

- تحلیل قطر مردمک فقط برای سطح Spread Text انجام شد.

- نبود تحلیل آماری (مانند تست معنی‌داری) تفسیر را محدود کرد.

۶ نتیجه‌گیری (Conclusion)

این گزارش نشان داد که با استفاده از ردیاب چشم Pupil Labs می‌توان مشارکت بصری دانشجو در یادگیری الکترونیکی را به صورت کمی سنجید. شاخص مشارکت مبتنی بر فیکسیسیون‌ها، قطر مردمک و سرعت نگاه، تفاوت‌های توجه به محتوای بصری را آشکار کرد. این روش می‌تواند برای طراحی سیستم‌های آموزشی تطبیقی استفاده شود تا محتوا بر اساس سطح درگیری فراگیر تنظیم گردد.

۷ مراجع (References)

- Pupil Labs. (2023). Pupil Labs Documentation: Core and Invisible. Retrieved from <https://docs.pupil-labs.com/> [1].
- Sample Recording. (2023). Pupil Labs Sample Recording v2. Google Drive. Retrieved from https://drive.google.com/file/d/1vzjZkjoI8kESw8lBnsa_k_8hXPf3fMMC/view [2].