GRANT SUBMISSION PREVIEW: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

Edit The Regents of the University of California

Grant Submission ID: **834409**

Proposal Category: **Grant Submission**

Grant Submission Status: **Complete**

Summary and Detailed Narrative

Dollar Amount of this Funding Request:

175,172

Geographic Area Served by Grant Request (i.e. City, State, Region, National, International, online, etc):

International

Summary

EyeSwipe is a proposed fast eye-based user interface for digital devices, initially enabling ALS and other neurologically constrained patients to speak using their eyes much faster than now possible. The speed of human speech is typically 120-150 WPM. Fast eye-typists, with high-quality eye-tracking systems such as those from Tobii/Dynavox, rarely exceed ~20-25 WPM, even with auto-completion. EyeSwipe aims to make comfortable 60-80 WPM eye-typing feasible in 2020 by applying neural networks trained on videos of eyes typing known phrases, and eschewing dwell- and visual-feedback-related delays in the UI. The long-term project aim is to create a new *de facto* standard for eye-based control of digital connected devices, to be integrated into all operating systems (e.g., Linux, Windows, MacOS, Android, and iOS).

Problem

Digital connected devices, now part of the quotidien for billions of people worldwide, provide user interface adaptations for users with visual, hearing, and motor disabilities. However, a significant population can neither speak nor use a touch-based interface effectively. Most persons with ALS (aka PALS) will need eye-based communication aids at some stage. Current eye-based typing systems such as those from Tobii/Dynavox are based on high-precision eye tracking systems. These special-purpose eye-tracking hardware and purpose-built software systems are not standardized and are costly, at $5,000-15,000. Moreover, these systems all give visual feedback on a letter-by-letter basis, which is known to slow the interaction. There is a unique need for fast, eye-based typing for PALS to communicate using off-the-shelf devices, and an opportunity to fulfill this need thanks to powerful phones and tablets combined with neural network based machine learning algorithms.

Solution

As anticipated by Marggraff, et. al. in the published US patent application 2017/0123492A1, we plan to build a fast eye-tracking analogue to a "swipe keyboard." We will make it available initially as an iOS or Android app on a smartphone or tablet to be mounted in front of a user's eyes.

It will have an eye-typing interface whose use consists of moving one's eyes as fast as comfortable from letter to letter, where the screen provides NO visible feedback as the user's gaze moves to each letter. Audio feedback may be provided at a word or phrase level, "whispered" to the user. Words are collected, and the sentence may be reviewed by the user upon completion of the entire sentence.

The system will not use traditional eye-tracking; rather, it will use AI in the form of an LSTM (Long Short Term Memory) RNN (Recurrent Neural Network) that has been trained to recognize unique gaze paths per word or phrase. It is trained by hundreds of hours of people who are given words to "type" using this configuration, and who produce as rapid-as-possible gaze paths for the words.

Hardware: Today's smartphones and tablets have substantial processing capability and a visible-light, high resolution front facing camera (e.g., 7 MP - iPad Pro, 10MP - Galaxy Note), enabling this new approach to eye-typing at a relatively low cost.
To enable use by PALS, once initial target devices have been identified, we will work with the CITIRIS fabrication and 3d-printing workshop to create rigs to enable fastening each device to a wheelchair optimized for comfortable eye-typing.

Implementation

The project is on a fast timeline because of the life-changing impact for the target population. Development follows three tracks: Data Collection, Model Building, and the Final Keyboard App. A fourth track integrates PALS accessibility issues through the other three.

Project plan:

Month	Track 1: Data collection	Track 2: Model building	Track 3: UX/App development	Accessibility
1	Build Data Collection App (2 months). Implement keyboard, video capture/labeling/storage.	Using fake data, try 5 RNN architectures	Study existing interfaces Tobii/DynaVox, begin prototyping, Create a slide deck with design prototype to present to target population	Contact target population, PALS - UCSF medical center, UC Davis Med Center, John Costello at Harvard, Set up first meeting to discuss UX
2	--	--	Iterate on design - needs start/stop commands, cropping, eye detection, keyboard	Work with PALS, iterate
3	App created, data collection begins: outreach, training, QA.	--	Iterate on design - needs start/stop commands, cropping, eye detection, keyboard	--
4	**8/15/20:** Data collection completed, ready for model development	Train with collected data	Fake demo : type a sentence, see it appear (model not ready yet)	--
5		**9/15/20**: RNN model built, ready to integrate with app	Integrate Model	--
6			**10/15/20**: Testing, app ready	Testing with target population

We will conduct Data Collection through a custom-built app. Using a device's front-facing camera and an on-screen keyboard, we will collect videos of a user's eyes as they eye-type, or "swipe" a prescribed phrase or sentence. These videos, ideally created by target users, are essential to the creation of an EyeSwipe dataset to train the neural network model to function in the field.

We create a corpus of ground-truth phrases combining snippets of everyday speech and conversation - atypical of NLP corpora - as well as snippets of typed exchanges such as email and documents more typical of NLP corpora widely available. In creating the everyday speech corpus, to reflect the practicality and informality of speech, we have begun scraping popular YouTube subtitles against a lexicon of common english words. Grooming the corpora will take 4 weeks, with 2 engineers at half-time.

Work has begun on the Data Collection app. An undergraduate intern, Max Sharnoff, generated an initial corpus of phrases using the method described, and began work on a server-client system by which phrases would be selected and transferred to the app for a user to eye-type. His preliminary work is publically accessible here: Github EyeSwipe repository. The goal of this initial app is to enable fast, high-quality, distributed data collection, which will facilitate building NN models and experimenting with a variety of architectures and design choices.

The second development track, RNN Model Building, aims to build a LSTN RNN model that achieves a target baseline error rate below 5% when trained on the dataset created from the first track. This stage will require 2 additional engineers from among graduate and

undergraduate students studying applied machine learning, with experience in neural networks, preferably LSTM's on video. Work can begin with a "stub" data set for training created by the engineers themselves to iterate and test various network architectures, among others the LSTN's used for lip-reading.

The third development track is that of the UI/UX, creating an EyeSwipe keyboard mobile app. Iteration and experimentation in UI and experience design is critical for an app tuned to the task of comfortable, high-speed eye-typing by the target population. The 4 engineers will work together with a UI/UX expert, to be engaged early in the project and work through a typical UX design process.
A fourth track in the project plan engages an accessibility expert, both as a stand-in for ALS users and a channel for regular iteration and testing by ALS and other neuro-constrained patients throughout the development process. We expect to collaborate with John Costello, Director of the Augmentative Communication Program at Boston Children's Hospital, who has daily interactions with the target population and is eager to support or work on this project.

The development and design of the app will be a continuous delivery process, as we work together with PALS and volunteers to best inform our choices in creating an appropriate user experience for the final app.

Note that, by using stub software and mock data sets, all tracks can proceed to some degree in parallel to compress the timeline. Our aim is to produce a proof-of-concept app that can be used by PALS within 6 months, allowing several months' lead time to reduce software defects and make the first "alpha" version of the complete system available to PALS before the holiday season in 2020. While this goal is ambitious, our team is inspired by the prospect of conversation at near-natural speeds with our PALS friends, and enabling the same for PALS around the world.

New or Existing Project

This is a new project originally conceived in 2015 by inventor and entrepreneur Jim Marggraff with his cofounders in the course of developing an eye-based interface for virtual reality as founding CEO of EyeFluence, Inc. Jim and his team laid out the core technical aspects of the project at that time and we are honored to have the opportunity to bring his ideas to fruition, with Jim's and several of his cofounders' blessing. As a member of the board of venerable ALS activist Steve Gleason's Foundation, Jim has offers not only his collaboration, but his personal and professional connections in taking this project forward. Their ideas for a fast eye-typing interface were published as part of a patent application filed August 15, 2016. Our broader aim is to create open-source software libraries that the major digital firms will integrate into their core OS accessibility suites. Last year, Jim connected with Mark Nitzberg, ED of the Center for Human-Compatible AI and part of the leadership of the Berkeley AI Research Lab. Over the past 8 months, Mark has supervised a volunteer intern exploring the data collection stage of the project and establishing the initial team.

Numbers Impacted

The initial target population for our project includes Persons with ALS (amyotrophic lateral sclerosis), or PALS. ALS affects as many as 30,000 in the US alone, with 5,000 new cases per year. Across the globe ALS is projected to reach 376,674 in 2040. Moreover, it is likely to be a breakthrough to a similar number of neurologically constrained users. It may also apply to mass-market and specialist uses in professionally constrained settings where noise prevents the use of voice-control and the user's hands are otherwise engaged, such as large machine operators and aircraft pilots. Ultimately, fast eye-based screen interaction is an essential part of a full-spectrum interactive VR or AR experience, whose market potential is in the hundreds of millions.

Metrics

We are committed to building a usable proof-of-concept app that enables comfortable, near-natural speed eye-typing for PALS, in the form of one or more apps for iOS and/or Android OS, whose components will be released as an open-source library. This library will be structured to maximize its attractiveness to serve as the basis for purveyors of digital operating systems (Linux, Windows, MacOS, Android and iOS) to integrate eye-swipe functionality into standard accessibility features.

Our success criteria are measures of (i) the speed at which a PALS can communicate (in words per minute), with a target of 40 WPM minimum, 50 WPM median and 60-80 WPM peak, (ii) the accuracy of the app against labeled test data, with an initial target of 95% accuracy, and (iii) a qualitative measure of comfortable use, with an initial target of "at least as comfortable to use as my current system when typing at comfortable, but lower, speeds".

Sustainability

Our aim is ambitious: to make fast eye-typing a global standard user interface - the very definition of sustainable. Although this grant request funds the project only to the point of creating a proof of concept eye-typing app, we are dedicated to using the project to attract and retain the big digital companies into a long-term partnership towards the creation of an international standard for eye-based universal access to digital devices. During the project period we plan to complete the larger project plan, with rough parameters expected to be a budget of $2-3M in cash and in kind from public and private sources, and a timeline of 3 years to creation and adoption by all relevant companies of the standard. Importantly, in the next phase of the project, we intend to develop simultaneous ports of the software libraries begun during this Phase to Linux, Windows, MacOS, iOS and Android in concert with the respective accessibility groups within Microsoft, Apple and Google, assuring that the resulting features are not captive to a single platform provider but truly universal standards.

Partnerships

At this time we have only begun the process of building links into numerous organizations as sponsors and collaborators. The low-hanging fruit connected to UC Berkeley and CITRIS as well as friendly organizations and agencies is only our starting point, and includes:

Initial Potential Funding Options:

1. Foundation sources including Team Gleason
2. CITRIS Seed Grants
3. CABHI (Toronto)
4. CITRIS Foundry
5. SBIR from NIA and other government agencies
6. Various individual donors
7. NIDILRR
8. DARPA
9. Microsoft, Google, Apple
10. VSP
11. Weill Neurohub

Initial Potential Collaborators within the CITRIS network:

1. Fung Fellows
2. UC Davis Health
3. UC Berkeley School of Public Health
4. Ram Akella/Gerald Friedland - CIBDA
5. CITRIS Invention Lab
6. CITRIS Digital Health Resource Center (UCD Health)

Sharing Results

EyeSwipe is an intentionally open-source project, with its home at the public university that holds an undisputed position as the top computer science program in the world for over a decade. Sharing our results is central to the purpose of the project. We will capture video testimonials of users in the target population of PALS using EyeSwipe, as well as publish the software and data corpora as open-source. We intend moreover to create friendly (if fierce) competitions and hackathons to encourage continuous improvement and creative and ever broader applications of this erstwhile arcane mode of human-machine interaction. In the longer term, we see ways to build out eye-based user experiences that go beyond typing, to enable full control of any device with a screen and front-facing camera to catalyze the adoption of EyeSwipe into all operating systems.

Video (optional)

Uploaded File(s)

Grant Submission Title: The Regents of the University of California

Project Budget

EyeSwipe: budget
foundation-budget-template-2020.xlsx
Six-month budget aiming to bring an app to families by the holiday season 2020.

Financial Statements

EyeSwipe: Statement of Financial Position
2018-19_UCB_financial (dragged).pdf
UC Berkeley Statement of Financial Position

EyeSwipe: Income Statement
2018-19_UCB_financial (dragged).pdf
UC Berkeley Statements of Revenues, Expenses, and Changes in Net Position

Attachments and Additional Information

EyeSwipe: Bios of key personnel
EyeSwipe bios.pdf
Brief bios of Mark Nitzberg, PhD (PI) and key consultants Jim Marggraff, Dr. José-Alain Sahel, David Stiehr, Peter Milford, and John Costello.

EyeSwipe: Budget Justification
Nitzberg CTAF Budget Justification.docx
Standard form budget justification provided to the UC Berkeley sponsored projects office.

EyeSwipe: Institutional letter and letters of support
EyeSwipe support letters.pdf
Letters of support: Institutional letter from UC Berkeley for PI Mark Nitzberg, and letters from John Costello (Boston Children's Hospital), Jim Marggraff (Marggraff LLC, project concept co-inventor), Jose-Alain Sahel (UPMC Eye Center), David Stiehr (formerly of Eyefluence), Peter Milford (formerly of EyeFluence).

Executive Summary

Federal Tax ID Number
94-6002123

Are you a 501(c)(3) nonprofit in good standing with the IRS?
Yes

Total Project Budget:
175,172

Organization's total budget:
2,100,000,000

Start Date of your Fiscal Year:

2019-07-01

Please mark all of the categories that apply to the populations served in this grant
Aging
Cognitive/Neurodiversity
Mobility

Please mark all of the categories that apply to the type of program covered in this grant
Combating Social Isolation
Create Awareness of Tech Solutions
Direct Service
New Technology Development
Pilot of Technology
Promote Accessible/Age Friendly Design
Research
Use of Advanced or Emerging Technology

Previous Grants

Has your organization previously received a grant from the CTA Foundation?
No

If you have received a previous grant, please give a brief update on the status of your most recent grant from the CTA Foundation: