HAL
open science

Interaction par le regard : évaluation du retour d'information progressif

van Bao Nguyen, Francis Jambon, Gaëlle Calvary

To cite this version:

van Bao Nguyen, Francis Jambon, Gaëlle Calvary. Interaction par le regard : évaluation du retour d'information progressif. IHM'14, 26e conférence francophone sur l'Interaction Homme-Machine, Oct 2014, Lille, France. ACM, pp.153-158, 2014. hal-01090436

HAL Id: hal-01090436

https://hal.science/hal-01090436v1

Submitted on 3 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Interaction par le regard : évaluation du retour d'information progressif

Van Bao NGUYEN
LIG, Université Grenoble-Alpes
BP53, 38041 Grenoble cedex 9, France
e-mail : baonv90@gmail.com

Francis JAMBON
LIG, Université Grenoble-Alpes
BP53, 38041 Grenoble cedex 9, France
e-mail : francis.jambon@imag.fr

Gaëlle CALVARY
LIG, Université Grenoble-Alpes
BP53, 38041 Grenoble cedex 9, France
e-mail : gaelle.calvary@imag.fr

© ACM, 2014. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Actes de la 26ième conférence francophone sur l'Interaction Homme-Machine, 2014.
http://dx.doi.org/10.1145/2670444.2670463

Résumé
L'utilisation de l'oculométrie pour l'interaction par le regard se caractérise, dans le cas d'une approche monomodale, par un fort couplage entre perception et action : il n'est pas possible, dans le cas général, de distinguer une prise d'information d'une action de l'utilisateur. C'est pourquoi nous proposons une nouvelle approche, basée sur un retour d'information progressif, permettant une réduction de ce couplage. Les premiers résultats expérimentaux sont encourageants et indiquent que l'interaction par le regard peut être une technique d'interaction crédible dans certains contextes d'usage. Ils mettent aussi en lumière l'intérêt du retour d'information progressif.

Mots clés choisis par les auteurs
Oculométrie, interaction par le regard, Midas Touch, retour d'information progressif.

Mot clés de la classification ACM
H.5.2. [Information Interfaces and Presentation]: User Interface---Input devices and strategies, Interaction styles; I.3.6. [Computer Graphics]: Methodology and Techniques---Interaction techniques.

Problématique

Notre cadre d'étude s'intéresse à l'utilisation de l'oculométrie pour l'interaction homme-machine dans les situations où l'utilisateur ne souhaite pas utiliser ses mains, que ce soit pour des raisons de sécurité (e.g. chirurgien souhaitant consulter des radios sans utiliser ses mains pour garantir la stérilité du champ opératoire), de praticité ou de confort (e.g. personne consultant une recette de cuisine pendant que ses mains sont occupées à réaliser la dite recette).

En nous adressant a priori à tous les utilisateurs, notre approche se distingue des travaux concernant l'assistance des personnes avec handicap moteur. Notamment, nous nous intéressons à des usages de l'interaction par le regard qui peuvent être ponctuels, en complément d'une utilisation plus classique comme la souris ou le tactile. Ainsi, notre approche se doit de ne pas imposer des modifications profondes du paradigme d'interaction initial. Dans ce contexte, l'approche proposée se base également sur une interaction ne nécessitant pas d'apprentissage a priori.

En interaction par le regard, l'approche la plus naïve consiste à utiliser la position du regard dans le repère de l'écran comme un substitut direct pour le pointeur et le bouton d'une souris. Or, le regard n'est pas naturellement utilisé par l'humain pour agir sur l'environnement et plus généralement transmettre de l'information. Il est plutôt sollicité pour percevoir de l'information (sauf exception en relation humain-humain), de plus, certains mouvements du regard peuvent être la conséquence d'une activité réflexe.

Dans cette approche naïve, tout changement de position du regard de l'utilisateur entraînera le pointeur de la souris à se déplacer, et chaque déplacement pourra lancer une nouvelle commande. Autrement dit, cette approche associe au regard une capacité d'action sur l'interface, et toute prise d'information pourra se voir interprétée comme une action. C'est le verrou du "Midas Touch" [3] inspiré du Mythe de Midas qui transformait en or tout ce qu'il touchait, et ainsi ne pouvait plus ni manger ni boire.

Ainsi, l'utilisation naïve du regard comme dispositif d'entrée provoque un très fort couplage entre la prise d'information et l'action. L'objectif de nos travaux est de permettre l'utilisation du regard comme mode d'interaction unique, tout en contournant le verrou du "Midas Touch". Autrement dit, il s'agit de déterminer un paradigme pour assurer le découplage entre la perception et l'action lorsque l'on souhaite utiliser l'oculométrie comme un moyen unique d'interaction.

État de l'art

Les recherches en oculométrie remontent à la fin du XIXe siècle lorsqu'en 1879, Louis Javal a découvert des mouvements saccadés au niveau de l'œil dans une tâche de lecture. En 1908, Edmund Huey [1] a construit le premier oculomètre permettant de suivre le mouvement des yeux d'une personne en train de lire. Dans les années 50, Alfred L. Yabus [9] a défini ces mouvements en termes de fixations et de saccades, concepts toujours utilisés actuellement.

L'utilisation de l'oculométrie pour l'interaction homme-machine n'a vraiment été étudiée qu'au cours de ces trente dernières années, pendant lesquelles nous pouvons identifier cinq étapes remarquables :

En **1982**, Richard A. Bolt [2] a été le pionnier en proposant plusieurs approches pour utiliser le mouvements des yeux en interaction homme-machine. L'une des approches proposées nécessitait de faire appel à la reconnaissance vocale pour désambigüiser certaines des commandes issues du regard.

En **1991**, Jacob et al. [3] ont utilisé l'oculométrie pour implémenter un système interactif comportant les principales interactions (sélection, déplacement, scroll, menu, etc.). Ils ont également identifié le problème du couplage perception/action et c'est à ces auteurs que l'on doit la dénomination de « Midas Touch ». Jacob et al. ont proposé plusieurs techniques d'interaction combinant le regard et l'utilisation du clavier, ou le regard seul par la technique du « dwell » (déclenchement lié à la durée de fixation). Leurs travaux ont posé les bases de l'interaction par le regard. À la suite de Jacob et al., d'autres chercheurs ont travaillé sur l'utilisation du clavier et de la souris pour éviter le « Midas Touch », comme Vetegaal [4] ou Kumar et al. [5].

En **1999**, Zhai [6] a introduit la technique de « Magic Pointing », dédiée à améliorer la précision du pointage à la souris. En « Magic Pointing », le pointeur est automatiquement déplacé vers la zone dans laquelle l'utilisateur regarde. Il s'agit d'une approche un peu différente des précédentes, car ici le regard intervient non pas comme principale technique d'interaction, mais comme assistance pour améliorer la performance.

En **2007**, Drewes et Schmidt [7] ont mis en œuvre une interaction visuelle qui prend en compte les positions que l'utilisateur regarde continuellement, puis cherche à classifier les « gaze gestures » à partir des mouvements du regard depuis cette position vers l'une des 8 directions de sortie possibles.

En **2013**, Biswas et al. [8] ont proposé une nouvelle forme d'interaction destinée aux handicapés moteurs en employant une variante du « blink » (déclenchement lié au clignement des yeux) pour réaliser des tâches comme la sélection d'objet ou l'activation de commande.

En synthèse, l'interaction par le regard a fait l'objet de nombreux travaux depuis plus de trente ans, sans pour autant permettre à cette technique d'interaction de se diffuser dans le grand public. En effet, il existe un consensus parmi les auteurs pour présenter le « Midas Touch » comme principal verrou de la technique. Les auteurs ont répondu à cette problématique via des approches multimodales (voix, clic souris, touche clavier), ou monomodales (« blink », « dwell », « gaze gestures »). Cependant, aucune de ces approches ne s'est généralisée.

Contribution

Une action de l'utilisateur dans un environnement d'interaction graphique de type WIMP peut généralement se scinder en deux informations : une position donnée, par exemple, par les coordonnées du curseur de la souris, et un déclencheur, qui peut être l'appui sur le bouton de cette même souris.

Si l'on cherche à déterminer ces deux informations en interaction monomodale par le regard, la position sur l'écran est déduite assez simplement par la position du regard. Concernant le déclencheur, la littérature propose notamment le « dwell » ou le « blink ». Nous avons écarté a priori le déclencheur « gaze gesture » car il implique l'apprentissage d'une grammaire de

gestes, ce qui n'est pas adapté à notre contexte d'usage ponctuel.

Le « dwell » consiste à déclencher une commande si l'utilisateur maintient la position de son regard fixe pendant un temps donné, dans une zone donnée de l'écran. Le « blink » consiste à lier le déclenchement à un clignement des yeux lorsque le regard de l'utilisateur est à l'intérieur d'une zone donnée de l'écran. Ces deux techniques ont leurs limites, et des déclenchements non désirés peuvent se produire très facilement lorsque l'utilisateur regarde trop longtemps un bouton de commande, ou s'il cligne des yeux de manière réflexe en regardant une commande.

Notre hypothèse est que le fonctionnement du déclencheur n'est pas assez explicite pour l'utilisateur. En effet, les retours d'information proposés dans la littérature se contentent d'informer l'utilisateur au moment du déclenchement de l'action, comme cela se produit lorsque l'on clique sur un bouton avec une souris. Nous supposons qu'il est nécessaire de ne pas se limiter à la seule indication du déclenchement, mais d'apporter un retour d'information à grain fin sur l'ensemble des étapes du processus d'interaction.

C'est pourquoi, nous proposons d'apporter un retour d'information de type progressif qui concrétise visuellement les étapes et le temps requis/restant pour activer une commande ou sélectionner un élément dans une liste. Nous avons appliqué ce principe aux deux types de déclencheurs « dwell » et au « blink ».

Application au déclencheur « dwell » (figure 1) : Ce déclencheur est bien adapté aux actions qui ne nécessitent pas une lecture de longue durée du label de l'action, car ainsi, le temps de lecture requis peut être inférieur au seuil de déclenchement. Ce déclencheur peut être utilisé, par exemple, pour activer un bouton. Dans notre implémentation, l'activation est déclenchée, si, une fois passé le seuil de sélection à T_{fix} (correspondant à la durée minimale d'une fixation), l'utilisateur maintient son regard sur l'objet jusqu'à T_{init}.

Un retour d'information progressif (sous forme de barre de progression située à l'intérieur du bouton) est utilisé entre T_{fix} et T_{init} de manière à permettre à l'utilisateur de visualiser le temps restant jusqu'à l'activation (sans quitter des yeux le bouton). L'utilisateur peut alors, en cours d'interaction, porter son regard en dehors de l'objet et ainsi ne pas déclencher l'action.

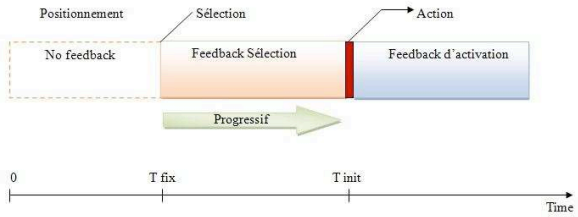

Figure 1. Déclencheur « dwell » avec retour d'information progressif.

Application au déclencheur « blink » (figure 2) : Ce déclencheur est bien adapté aux actions nécessitant une lecture de longue durée du label de l'action, car sa période d'activation peut être bornée. Il peut être utilisé, par exemple, pour sélectionner un élément dans une liste. Le principe de ce déclencheur est de ne permettre une activation par un clignement des yeux que pendant un intervalle de temps limité, afin d'éviter un déclenchement involontaire sur une activité réflexe. Dans notre implémentation, l'activation n'est possible qu'entre T_{init} et T_{exp}, et pendant cet intervalle de temps,

un retour d'information progressif (sous forme de décompte situé à l'intérieur du label) indique à l'utilisateur combien de temps il lui reste pour effectuer un clignement des yeux (sans quitter des yeux le label), et ainsi déclencher l'action (figure 2 : cas b) ou pas (figure 2 : cas a).

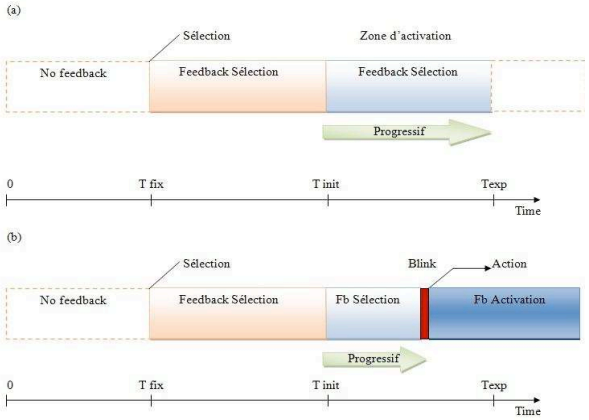

Figure 2. Déclencheur « blink » avec retour d'information progressif.

Nous avons mis en œuvre ces deux techniques d'interaction sur un démonstrateur : l'application EyeMail. Cette application est un logiciel simplifié de lecture de courriers électroniques, qui permet d'accomplir les principales tâches de ce type de logiciel : la sélection d'un courrier dans la boîte de réception, la composition d'un nouveau courrier, l'ajout d'une image jointe, son envoi, etc. Toutes ces commandes peuvent être réalisées en utilisant uniquement le regard. Nous avons ainsi employé les deux techniques d'interaction précédemment proposées :

- Activation par « dwell » : nous avons utilisé ce déclencheur pour activer les commandes (composer un e-mail, envoyer un e-mail ou attacher une image) ;
- Sélection par « blink » : nous avons utilisé ce déclencheur pour sélectionner un élément dans une liste (sélection des e-mails, sélection des images).

Évaluation

Nous avons réalisé des expérimentations utilisateur afin d'obtenir une première évaluation de l'apport effectif du retour d'information progressif. Nous avons testé les deux techniques d'interaction « dwell » et « blink », chacune avec et sans retour d'information progressif. Dans les deux cas, nous avons conservé le retour d'information présent lors du déclenchement de la commande. Pour ces expérimentations, nous avons utilisé trois logiciels : l'application EyeMail pour l'habituation, une application « Groupes de boutons » pour évaluer le « dwell » (le sujet doit activer le bouton indiqué dans la consigne parmi les 8 boutons disponibles), et une application « Liste d'éléments » pour évaluer le « blink » (le sujet doit sélectionner l'élément indiqué dans la consigne parmi une liste de 8 éléments disponibles).

Le protocole expérimental débutait par une calibration de l'oculomètre, puis le sujet pouvait s'habituer avec l'application EyeMail pendant le temps qu'il souhaitait. Ensuite, il effectuait successivement 4 tests différents : avec chacune des deux techniques, et avec/sans retour d'information progressif. L'ordre des techniques comme celui de la présence du retour d'information progressif étaient alternés. Les résultats ont été analysés avec les données suivantes : nombre d'activations réussies, nombre d'activations en dehors de la consigne, et avis subjectif des utilisateurs. En tout, 13 sujets,

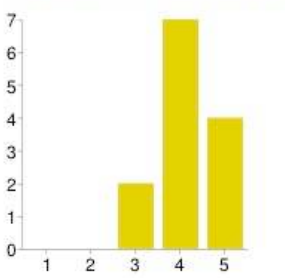

Figure 3. Avis des utilisateurs concernant la facilité d'utilisation du « dwell ».

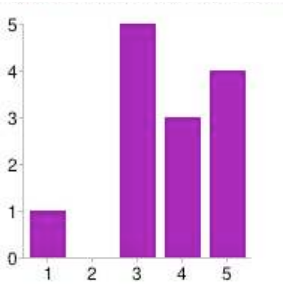

Figure 4. Avis des utilisateurs concernant la facilité d'utilisation du « blink ».

principalement des étudiants et des chercheurs en informatique, ont participé aux expérimentations.

Tout d'abord, concernant la technique du « dwell », 12 sujets sur 13 ont préféré la présence d'un retour d'information progressif, ils ont tous donné un avis neutre ou favorable à la facilité d'activation des commandes par cette technique (figure 3). La note moyenne est de 4,15 sur un barème de 1 à 5 (5 étant le plus favorable). Concernant le déclenchement par « blink », les sujets ont également majoritairement préféré le feedback progressif : 10 sujets sur 13. La note moyenne donnée par les utilisateurs a été de 3,7 (figure 4). Les résultats sont globalement moins positifs que pour la technique du « dwell ».

Enfin, en prenant en compte le nombre d'activations en dehors des cibles dans les deux modes (avec et sans retour d'information progressif), les interactions avec retour d'information progressif se montrent généralement plus efficaces que celles sans retour d'information : les activations en dehors de la cible sont au total de 4 avec retour d'information progressif et de 8 lorsque l'on ne l'utilise pas.

Conclusion & Perspectives

Bien que le nombre de sujets ne soit pas significatif et qu'ils ne soient pas représentatifs de la population générale, ces premiers résultats sont encourageants et indiquent que l'interaction par le regard est une technique crédible dans certains contextes d'usage, et que la présence d'un retour d'information progressif semble être préférée par les utilisateurs.

Ces premiers résultats devront cependant être confirmés par la mise en œuvre de nouvelles expérimentations comportant un panel plus important d'utilisateurs et en variant les conditions de l'expérimentation. Notamment, il serait nécessaire de développer des applications spécifiques à un domaine particulier (santé, domotique, etc.) et de les évaluer en situation écologique.

Références

[1] Huey, E. B. *The Psychology and Pedagogy of Reading* (Reprint). MIT Press (1968 originally published 1908).

[2] Bolt, R. A. Eyes at the interface. In proc. *CHI'82*. ACM (1982), 360-362.

[3] Jacob, R. J. K. What you look at is what you get: eye movement-based interaction techniques. In proc. *CHI'90*, ACM (1990), 11–18.

[4] Fono, D., and Vertegaal, R. EyeWindows: evaluation of eye-controlled zooming windows for focus selection. In proc. *CHI'05*, ACM (2005), 151–160.

[5] Kumar, M., Paepcke, A., and Winograd, T. EyePoint: practical pointing and selection using gaze and keyboard. In proc. *CHI'07*, ACM (2007), 421–430.

[6] Zhai, S., Morimoto, C., and Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. In proc. *CHI'99*, ACM (1999), 246–253.

[7] Drewes, H., and Schmidt, A. The MAGIC touch: Combining MAGIC-pointing with a touch-sensitive mouse. In proc. *INTERACT'09*, Springer-Verlag (2009), 415–428.

[8] Biswas, P. and Langdon, P. A new input system for disabled users involving eye gaze tracker and scanning interface. *Journal of Assistive Technologies*, vol. 5, iss. 2 (June 2011), 58–66.

[9] Yarbus, A.L. Eye Movements and Vision. New York: Plenum press (1967).