

Explorateur Audio-Tactile d'écran vidéo pour aveugles

C. Liard¹, A. Beghdadi²

¹ GAIB / CNAM 3 bld Pasteur 75015 PARIS, liard@cnam.fr

² L2TI -Institut Galilee, Université Paris 13, 93430 Villetaneuse, Beghdadi@galilee.univ-paris13.fr

Résumé

Aujourd'hui, le texte, le graphique et l'image sont intimement mêlés pour faciliter la compréhension de l'information à transmettre. Le dispositif pour personnes aveugles que nous présentons dans cet article est destiné à l'exploration de graphiques ou d'images présentés sur un écran d'ordinateur. Il se compose d'un ensemble d'outils sonores et tactiles que l'utilisateur peut activer à tout moment pour extraire des graphiques et, dans certains cas, des images qui accompagnent un texte, différentes informations nécessaires à sa compréhension et à son utilisation. C'est un dispositif dont la vocation est de compléter les systèmes de lecture de textes en braille ou par synthèse de parole.

I - Introduction

L'accès à l'information qu'elle soit présentée sous la forme de texte, de graphique ou d'image constitue pour les personnes déficientes visuelles ou aveugles une difficulté permanente dans leur vie quotidienne scolaire ou professionnelle. Aujourd'hui, les nouveaux moyens de communication utilisent couramment ces multiples formes de représentation de l'information. Internet et les logiciels modernes en font un usage intensif. Des solutions existent pour accéder au texte, notamment avec les terminaux Braille et les synthétiseurs de parole. Par contre, il n'y a pas de solution satisfaisante pour consulter, de façon interactive, des informations présentées sous la forme de graphiques ou d'images. Cet obstacle difficilement franchissable pour les personnes aveugles risque à terme de les exclure d'un champ d'informations qui s'impose de plus en plus comme le complément indispensable du texte. Le système que nous présentons ici a pour objectif de permettre à un non-voyant de consulter des informations texte, graphique ou image présentées sur un écran d'ordinateur sans perdre les attributs, comme la mise en page ou la couleur. Nous tentons d'apporter une première ébauche de solution à ce difficile problème. Ce système exploite les ressources de l'espace sonore et tactile. Nous utilisons notamment des ressources encore inexploitées comme la localisation spatiale de Sources Sonores Virtuelles (SSV).

II - Situation du problème

Lorsqu'une personne déficiente visuelle est en relation avec un ordinateur et plus particulièrement avec un logiciel, elle doit d'une part localiser la position du curseur souris à l'écran et se faire une image mentale de la disposition des éléments présentés à l'écran. Cette image doit être aussi proche que possible de la réalité afin d'éviter des erreurs qui peuvent conduire à des situations incontrôlables et souvent déprimantes pour l'utilisateur. Une bonne représentation des informations présentées à l'écran simplifie et améliore la qualité des échanges en milieu scolaire et professionnel ce qui a pour effet de faciliter l'insertion sociale de la personne déficiente visuelle.

III - Description Globale du système

Ce système repose sur l'utilisation conjointe du sens du toucher et de l'audition. La configuration d'ensemble de ce système est présentée sur la figure 1. Il se compose d'un programme situé dans la barre des tâches dans l'environnement Windows 98. L'utilisateur peut l'appeler à tout moment pour explorer le contenu de l'écran. Ce programme contrôle les éléments suivants :

- Une plage tactile vibrante (Optacon II) connectée à l'ordinateur par l'intermédiaire d'une liaison série. Elle est constituée par une matrice de 5 x 20 points située sous le doigt de l'utilisateur. Cette matrice donne à l'utilisateur une image binarisée en noir et blanc d'une fenêtre d'observation située sur l'écran. Dans certaine situation, cette plage tactile peut être associée à l'intégralité de l'écran. Dans ce cas, la zone de l'écran où se trouve le curseur souris est indiquée à l'utilisateur par l'activation d'un pixel vibrant de la plage tactile.

- Une carte son stéréophonique installée sur le bus de l'ordinateur et connectée à deux haut-parleurs. L'ensemble formé par les deux haut-parleurs et la tête de l'auditeur formant un triangle équilatéral de 1 mètre de coté dans le plan horizontal. L'exploitation de cet environnement sonore permet à l'utilisateur de localiser la position du curseur souris à l'écran avec une bonne précision. Ce résultat est obtenu en associant aux déplacements X et Y de la souris, une Source Sonore

Virtuelle (SSV) en mouvement. Une description sonore de l'écran est aussi disponible. Elle repose sur la transposition, dans l'espace sonore, des trois composantes couleur RVB pointées par la souris à l'aide d'un mélange de trois fréquences musicales. L'utilisateur peut ainsi connaître rapidement la composition et la disposition des éléments présentés à l'écran. Il peut aussi, si nécessaire, obtenir par synthèse de parole des informations plus précises sur la couleur qu'il pointe à l'écran.

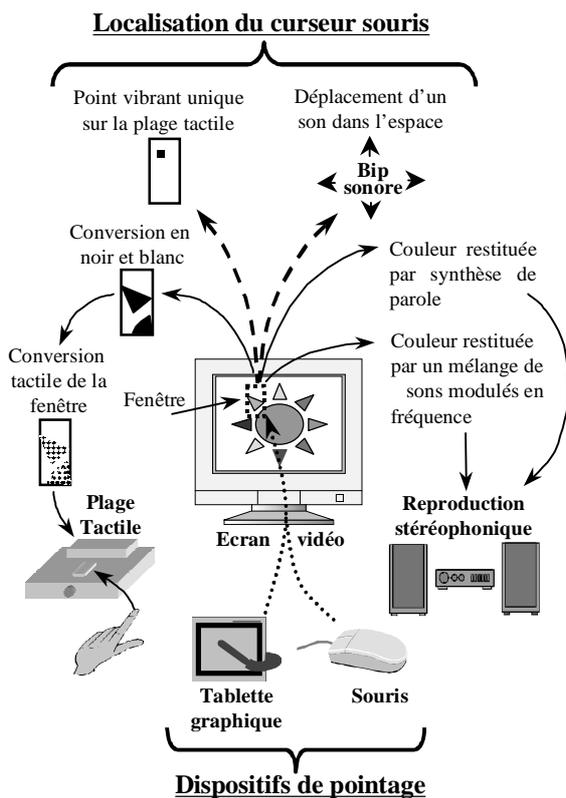


Figure-1 Configuration matérielle de l'explorateur et ressources Sonore et Tactile mises à la disposition de l'utilisateur pour convertir les informations contenues dans la fenêtre virtuelle.

IV – La souris acoustique et tactile

Se repérer rapidement dans l'écran lorsque l'information est composée à la fois de textes, de graphiques et d'images est un exercice difficile pour la personne aveugle. De plus, la disposition des éléments d'informations présentés à l'écran est très souvent associée à une signification qu'il ne faut sous estimer. C'est pour cette raison que nous proposons à l'utilisateur deux moyens originaux de se repérer dans l'écran qu'il peut utiliser indépendamment ou simultanément.

- Le feed-back acoustique :

Il y a dans l'espace sonore des ressources qui sont très peu exploitées comme l'audition directionnelle de sources sonores virtuelles. Cette caractéristique de l'espace sonore, totalement comparable à l'espace visuel par son aspect spatiale, est parfaitement adaptée pour restituer les mouvements de la souris avec une précision acceptable. C'est cette dimension de l'espace sonore que nous avons exploitée dans notre application pour restituer à l'utilisateur, d'une façon naturelle et sans apprentissage particulier, la position du curseur souris à l'écran. Pour cela, nous avons d'abord sélectionné une source sonore qui soit perçue par les auditeurs comme ponctuelle et peu fatigante. Ensuite, pour déplacer cette source sonore dans l'espace, nous avons combiné simultanément deux phénomènes ; la stéréophonie pour les déplacements horizontaux et la variation de fréquence du son pour les déplacements en hauteur. En effet, sur ce dernier point, nous avons exploité un phénomène psycho-acoustique qui conduit naturellement un auditeur à localiser les sons les plus graves vers le bas et les plus aigus vers le haut. La forme d'onde du signal utilisé pour restituer les mouvements du curseur souris est présentée sur la Figure 2.

Dans ces conditions, la position de l'index sonore virtuel dans l'espace frontal de l'auditeur est contrôlé par les mouvements en X et Y de la souris. L'utilisateur peut ainsi se rendre compte sans effort particulier de la position du curseur à l'écran.

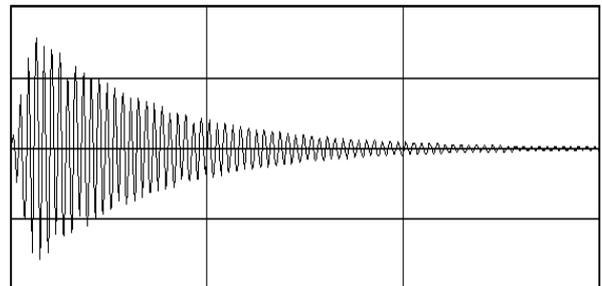


Figure-2 Le signal utilisé est sinusoïdal, il est modulé par une enveloppe dont la montée est linéaire et la décroissance exponentielle. La durée totale du phénomène est comprise entre 12 et 15 mS suivant la fréquence utilisée.

- Le feed-back tactile :

Malgré la faible résolution de la plage tactile, elle apporte une aide fiable à l'utilisateur pour localiser la zone de l'écran occupée par le pointeur souris. Dans ce cas, la totalité de l'écran est assignée à la plage tactile tout entière avec seulement un pixel vibrant actif indiquant la zone de l'écran où se situe le pointeur souris. L'utilisation conjointe des feed-back acoustique

et tactile permet à l'utilisateur d'obtenir la précision et la fiabilité nécessaire sur la position du pointeur souris à l'écran.

V – Exploration de l'écran

L'utilisateur dispose de deux méthodes pour explorer les informations texte, graphique et image présentées sur l'écran de l'ordinateur. La première est tactile et la seconde auditive. Ces deux méthodes ont leurs avantages et leurs inconvénients et peuvent se compléter avantageusement pour obtenir une meilleure perception des informations présentées à l'écran.

- Exploration tactile :

L'utilisateur peut découvrir les informations images présentées à l'écran à l'aide de l'Optacon. En effet, cet appareil fourni à l'utilisateur une représentation tactile binaire du contenu d'une fenêtre située à l'écran dont la position est contrôlée par les mouvements du curseur souris. Dans ces conditions, pour un utilisateur droitier, celui-ci contrôle, à l'aide de la souris, la position de la fenêtre de sa main droite et dispose sous l'index de sa main gauche d'une image tactile binarisée d'une portion élémentaire de l'écran. L'utilisateur peut donc se faire une représentation mentale de la totalité de l'information en explorant de proche en proche le contenu de l'écran. Dans certaines situations, pour faciliter l'exploration de l'information présentée à l'écran, l'utilisateur peut fixer à tout moment l'une des composantes X ou Y de la souris pour consulter l'image soit horizontalement, soit verticalement sans avoir à utiliser de guide mécanique.

- Exploration Sonore :

La représentation tactile binaire obtenue à partir de l'Optacon est beaucoup trop réductrice pour que l'utilisateur puisse se faire une représentation mentale satisfaisante du contenu de l'écran. D'autres ressources sont donc nécessaires à mettre en œuvre pour guider l'utilisateur vers une meilleure appréciation de l'information présentée à l'écran. Pour cela, nous avons associé à un pixel de l'écran une fenêtre élémentaire constituée du pixel concerné et de ceux qui l'entourent, soit au total 9 pixels. L'intérêt de cette fenêtre est de pouvoir évaluer certaines couleurs dont le degré de clarté est obtenu par une pigmentation organisée de pixels blancs. Cette fenêtre ainsi définie est alors réduite à un pixel fictif unique que l'on associe à un mélange de trois fréquences sonores pour former un son unique qui est restitué à l'utilisateur en stéréophonie. Ce principe d'exploration permet à l'utilisateur, avec un peu d'expérience, d'apprécier les objets pointés ainsi que leurs positions en exploitant les ressources de l'espace sonore.

La méthode que nous utilisons est la suivante : On considère la moyenne des trois composantes couleur R_i , V_i et B_i , des pixels P_i constituant la fenêtre que l'on associe ensuite à un pixel fictif P_f unique de composantes fictives R_f , V_f , B_f , tel que :

$$R_f = \sum_{i=1}^9 R_i / 9 \quad V_f = \sum_{i=1}^9 V_i / 9 \quad B_f = \sum_{i=1}^9 B_i / 9$$

Ensuite, les trois composantes de ce pixel fictif P_f sont associées respectivement à trois signaux sinusoïdaux S_R , S_V , S_B d'amplitude unitaire mais de fréquences respectives, F_R , F_V , F_B . Dans ces conditions nous avons :

$$S_R(t) = \sin(2\pi F_R t)$$

$$S_V(t) = \sin(2\pi F_V t)$$

$$S_B(t) = \sin(2\pi F_B t)$$

Les trois fréquences sonores F_R , F_V , F_B , sont respectivement modulées, par les trois composantes couleur correspondantes R_f , V_f et B_f , du pixel fictif P_f suivant les relations suivantes :

$$F_R = F_0 (\sqrt[12]{2})^{[12R_f/255]}$$

$$F_V = 2F_0 (\sqrt[12]{2})^{[12V_f/255]}$$

$$F_B = 4F_0 (\sqrt[12]{2})^{[12B_f/255]} \quad \text{Où [...] = partie entière}$$

Chaque composante de couleur est ainsi représentée par 12 notes d'une octave de la gamme musicale tempérée calculée à partir d'une fréquence de référence F_0 . Le signal sonore S restitué en stéréophonie et perçu par l'auditeur est alors donné par l'expression suivante :

$$S(t) = A(S_R(t) + S_V(t) + S_B(t))$$

Où A est un coefficient d'amplification. Dans ces conditions, l'auditeur peut apprécier rapidement les zones colorées ainsi que certains objets courants présentés à l'écran et d'en connaître leur disposition.

Dans le cas où l'utilisateur a besoin d'information plus précise sur la couleur qu'il pointe, il peut à tout moment faire appel aux fonctions de traduction de la couleur par synthèse de parole. Pour cela, nous avons défini 18 couleurs courantes

auxquelles nous ajoutons des adjectifs tel que « clair », « très clair » ou « très très clair » pour indiquer le degré de clarté de la couleur et d'une façon comparable les adjectifs « foncé », « très foncé » ou « très très foncé » pour indiquer la tendance sombre de la couleur.

VII – Discussion

Aujourd'hui, l'information ne se limite plus au texte mais devient un ensemble complexe d'éléments d'informations constitués à la fois de textes, de graphiques, d'images et même de séquences vidéos. Internet est un exemple frappant d'exploitation de l'ensemble de ces ressources. Le système que nous proposons a pour objectif l'exploration interactive de graphiques et d'images présentés à l'écran. Ce dispositif est composé d'un certain nombre de ressources sonores et tactiles qui permettent au non-voyant, avec un peu d'expérience et de motivation, de se faire une représentation plus ou moins précise des informations présentées à l'écran. L'aspect global du graphique ou de l'image est obtenu à l'aide des ressources acoustiques alors que les détails sont accessibles par des moyens tactiles. Si l'accès à l'image reste pour le moment difficile à appréhender, en revanche l'accès au graphique donne des résultats encourageants. La véritable vocation de ce système est surtout d'apporter une aide destinée à compléter les informations textes obtenues à partir d'un dispositif de lecture en braille ou par synthèse de parole, avec des éléments d'informations extraits des graphiques ou des images qui accompagnent un texte. Même si cette aide reste rudimentaire par rapport à l'ampleur des problèmes à résoudre, il n'en reste pas moins qu'elle apporte une aide importante au non-voyant qu'il s'agit de d'améliorer et de compléter par d'autres ressources.

Conclusion

Grâce aux ressources sonores et tactiles de cet explorateur d'écran, l'utilisateur peut apprécier les différents éléments d'informations, graphiques ou images, présentés à l'écran et élaborer sa propre représentation mentale. Pour cela, l'utilisateur dispose d'un ensemble de retours acoustique et tactile pour évaluer ; d'une part, la position du curseur souris à l'écran et d'autre part, le contenu de la zone écran qu'il pointe. L'utilisation d'un signal sonore stéréophonique constitué par un mélange de 3 fréquences est certainement une approche intéressante pour explorer la globalité d'un graphique ou d'une image en couleur. La traduction de la couleur par synthèse de parole apporte un complément indispensable et fiable permettant de peaufiner sa représentation mentale. Ce système possède un certain nombre de ressources pour accéder

au graphique et à l'image qu'il s'agit maintenant d'évaluer sur un petit groupe d'utilisateurs.

VIII – Références

1. Meijer, P.B.L., "An Experimental System for Auditory Image Representations" IEEE Transactions on Biomedical Engineering, Vol. 39, N° 2, p112-121, Feb 1992.
2. C. Capelle, C. Trullemans, P. Arno and C. Veraart, "A Real-Time Experimental Prototype for Enhancement of Vision Rehabilitation Using Auditory Substitution," IEEE Trans. Biom. Eng., vol. BME-45, pp. 1279-1293, Oct. 1998.
3. C. G. Martins and R. M. Rangayyan, "Experimental Evaluation of Auditory Display and Sonification of Textured Images," International Conference on Auditory Display (ICAD), Palo Alto, CA, November 2-5, 1997
4. Liard, C. (1995) "Représentation d'objets graphiques destinée à des personnes déficientes visuelles ou aveugles, reposant sur la localisation auditive spatiale de sources sonores virtuelles. Approches Psychoacoustique". Thèse, Université Paris VI, France.
5. G.Sainarayanan, R. Nagarajan, Sazali Yaacob, "Incorporating Certain Human Vision Properties In Vision Substitution by Stereo Acoustic Transform", Proc. IEEE, Vol. 1, pp60-63, ISSPA-2001, Kuala-Lumpur, Malaysia, 13-16 August 2001
6. C. Liard, A. Beghdadi, "An Audio Display Tool : The Sound Screen System", Proc. IEEE, Vol. 1, pp198-201, ISSPA-2001, Kuala-Lumpur, Malaysia, 13-16 August 2001.
7. T. Maucher, K. Meier, J. Schemmel, "An Interactive Tactile Graphics Display", Proc. IEEE, Vol. 1, pp. 190-193, ISSPA-2001, Kuala-Lumpur, Malaysia, 13-16 August 2001.
8. Sakamoto N, Gotoh T, Kogure T, Shimbo M, Clegg A., (1981) – Controlling sound-image localization in stereophonic reproduction. J Audio Eng Soc 29:794-799.
9. Blauert J, (1983) "Spacial hearing. The psychophysics of human sound localization. The MIT Press.