



HAL
open science

Augmentation d'un tableau blanc par des techniques de Vision par Ordinateur

François Bérard, Alex Pentland, Yann Laurillau

► **To cite this version:**

François Bérard, Alex Pentland, Yann Laurillau. Augmentation d'un tableau blanc par des techniques de Vision par Ordinateur. Conférence ASTI'01, 2001. hal-03286472

HAL Id: hal-03286472

<https://hal.science/hal-03286472>

Submitted on 14 Jul 2021

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.

Augmentation d'un tableau blanc par des techniques de Vision par Ordinateur

François Bérard, Sandy Pentland
MIT Media Laboratory
20, Ames street
Cambridge, MA 02139 USA
+1 617 253 0381
{ francois,sandy } @media.mit.edu

Yann Laurillau
Laboratoire CLIPS-IMAG
BP 53
38041 Grenoble Cedex 9, France
+33 476 514 489
Yann.Laurillau@imag.fr

RÉSUMÉ

Cette vidéo illustre le système "Tableau Magique". Le système est fondé sur un tableau blanc banal ou une table sur laquelle on écrit avec des feutres à encre effaçable. Une caméra et un projecteur vidéo sont orientés vers la surface. La caméra permet de capturer les inscriptions portées sur la surface, le projecteur permet d'afficher une version électronique des inscriptions sur la surface. Les utilisateurs interagissent avec les inscriptions électroniques par l'intermédiaires de jetons en plastique rouge. Le système permet de copier, déplacer, changer la taille et faire tourner les inscriptions. Les manipulations se font à deux mains et il est possible d'interagir à plusieurs utilisateurs simultanément.

Mot-clés

Réalité augmentée, tableau blanc, vision par ordinateur.

INTRODUCTION

Les tableaux blancs ou les tableaux à feuilles (flipchart) sont des outils largement utilisés car ils sont un support adaptés aux réunions informelles d'un petit groupe de personnes. L'accès à ces outils est facile et immédiat : tout le monde sait comment écrire avec un stylo et effacer les erreurs avec son doigt ou avec un effaceur. De plus, ces outils permettent à plusieurs personnes de travailler en même temps sur la même surface.

Cependant, la difficulté de capturer et d'archiver l'information portée sur un tableau blanc entraîne souvent sa perte. L'utilisation d'un tableau à feuille apporte une solution à ce problème mais la surface d'un tableau à feuille bien inférieure à celle d'un tableau blanc. Dans les deux cas, les surfaces inscriptibles sont physique et de fait elles ne permettent pas un archivage et une diffusion de l'information aussi facilement que dans le cas d'informations électroniques.

Pour répondre à ces problèmes plusieurs tableaux électroniques ont été conçus. Ces tableaux utilisent des dispositifs physiques permettant d'enregistrer la trajectoire et la couleur des stylos utilisés. Les informations enregistrées représente une copie électronique des informations portées sur la surface. Cependant les technologies de capture utilisées ont des limites :

- Seule la trajectoire du trait est capturée. La largeur du trait et la calligraphie sont perdues.
- La résolution de capture est parfois limitée. Une résolution spatiale insuffisante ne permet pas l'écriture de

petit symboles. Une résolution temporelle insuffisante ne permet pas l'écriture rapide.

- Enfin, il est nécessaire d'appliquer une pression sur le stylo supérieure à la pression utilisée avec un feutre normal (sous peine de ne pas enregistrer la trajectoire). C'est à nos yeux la limitation la plus gênante car elle impose une modification de l'ergonomie de l'écriture.

L'APPROCHE DE CAPTURE VISUELLE

Une autre approche pour numériser l'information portée sur le tableau est d'utiliser un tableau blanc normal et de mettre en oeuvre des techniques de vision par ordinateur. Une caméra est orientée vers la surface d'écriture afin de capturer l'information. C'est cette approche qui est utilisée dans les projets ZombieBoard [3] et le BrightBoard [5]. Le principal avantage est que la surface et les outils (stylos, effaceurs, doigts nus) ne sont pas modifiés. Ainsi, l'ergonomie de l'écriture n'est absolument pas modifiée. De plus, la capture visuelle permet une reproduction plus fidèle: la capture ayant lieu *après* l'exécution du dessin, il n'y a pas de problème de résolution temporelle. L'utilisation d'un appareil photo haute définition ou d'une caméra vidéo pouvant s'orienter et zoomer sur une petite zone de la surface assure une haute définition spatiale de capture. Enfin, la texture, la couleur, l'épaisseur des traits sont capturées par une capture visuelle.

La capture visuelle n'est pas idéale pour toute les tâches. La capture de la trajectoire du stylo facilite grandement la reconnaissance de l'écriture. L'approche de capture visuelle implique la mise en oeuvre de techniques de reconnaissance de caractères optiques, techniques plus coûteuse en calculs et dont le taux de reconnaissance est inférieure. La capture de la trajectoire permet également la reconnaissance en temps réel de certains gestes qu'on peut associer à des commandes telles que "couper" et "coller". Flatland [2] par exemple utilise la reconnaissance de gestes pour offrir un grand nombre de fonctionnalités à ses utilisateurs.

Nous choisissons l'approche visuelle car notre priorité est de conserver l'ergonomie de l'écriture sur tableau blanc. Nous pensons que le faible essor des tableaux électroniques commerciaux actuels est la conséquence de la modification imposée sur la façon d'écrire. Toutefois, nous désirons permettre l'activation de commandes grâce aux gestes. Nous proposons l'utilisation de jetons dont la trajectoire est

analysée par le système. En manipulant ces jetons, l'utilisateur peut envoyer des commandes au système.

NOTRE SOLUTION

Dispositifs physique

Notre système est composé d'une surface horizontale ou verticale (une table ou un tableau). La surface est un plan de couleur blanche sur lequel les utilisateurs écrivent et dessinent avec des feutres à encre effaçable.

Une caméra orientable, à facteur de zoom variable, est installée en face de la surface pour permettre la capture des inscriptions. Cet arrangement présente un problème lorsque l'utilisateur se présente devant la caméra et occulte une partie de la surface. Ce problème est observé principalement dans le cas du tableau. Sur la table, il est rare que les utilisateurs se penchent assez pour occulter la surface.

Un projecteur vidéo est installé près de la caméra. Il sert à la fois à projeter l'information capturée sur la surface et à projeter les retours d'informations du système (par exemple un cadre de sélection).

Les utilisateurs envoient des commandes au système par l'intermédiaire de jetons. Nos jetons sont de petits disques de plastique rouges de 4 cm de diamètre environ. Dans le cas du tableau, les jetons sont collés à de petits aimants pour qu'ils tiennent en place sur la surface verticale.

Lorsque la caméra n'est pas utilisée pour la capture des inscriptions, elle fournit un flux vidéo à un système de vision par ordinateur dont le rôle est de suivre la position de tous les jetons présent dans le champ de vision de la caméra.

Système de vision par ordinateur

Nous employons deux systèmes différents. L'un est dédié à la capture de bonne qualité des inscriptions. Ce système constitue une mosaïque des images obtenues en orientant la caméra vidéo et en zoomant vers les zones à capturer. Grâce à la mosaïque nous obtenons des images de haute définition de la surface (par exemple 3600x2700 points) à partir d'une caméra vidéo standard (640x480 points).

Le second système de vision par ordinateur suit en temps réel la position des jetons colorés sur la surface. Le système est initialisé en lui "montrant" un ensemble de pixels qui représentent un des objets colorés qu'il s'agit de suivre. Un ensemble de statistiques sont calculées sur la valeur de ces pixels. Ces statistiques sont utilisées pendant le suivi pour localiser les pixels du flux vidéo qui ressemble aux pixels montés lors de l'initialisation. Cette technique nous permet de suivre un grand nombre de jetons (jusqu'à 20 jetons la perte de performance du système est négligeable).

Interaction

Notre système offre la possibilité d'effectuer 3 types d'actions que nous détaillons maintenant.

- *Capture par sélection.* Pour sélectionner une zone, deux jetons sont mis en contact comme illustré sur les figures 1a et 1b. Les jetons sont ensuite séparés. Le système capture la zone située entre les deux jetons lorsque ceux-ci restent stables pendant 3 secondes (figure 1c). La sélection peut être annulée avant la capture en cachant

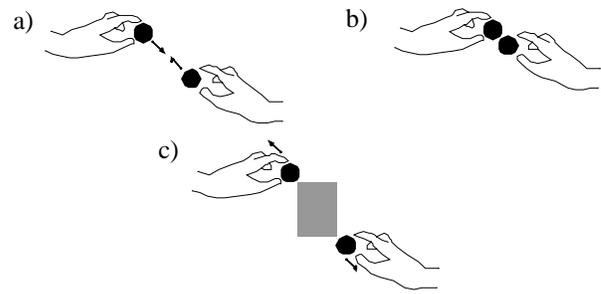


Figure 1 : Sélection d'une zone. Deux jetons sont mis en contact (a et b). Une boîte de sélection grise est "attachée" aux jetons (c). L'action de sélection se termine lorsque les jetons sont immobiles pendant 3 secondes.

l'un des jetons avec la main. Si la capture a lieu, la zone capturée (un "patch") est alors projeté sur la surface.

- *Déplacement / rotation / changement de taille d'un patch.* Lorsqu'une zone a été capturée, les deux jetons qui ont servis à définir la sélection deviennent des "poignées" sur le patch capturé: déplacer les poignées provoque le déplacement, la rotation, et le changement de taille du patch de façon à ce que la position relative des poignées par rapport au patch reste constante. Une poignée est détachée de son patch en la cachant avec la main. Un jeton qui n'est pas encore attaché peut être accroché à un patch en le déplaçant sur la surface du patch (avec un maximum de deux poignées par patch).
- *Destruction d'un patch.* Un patch est détruit en réduisant sa taille à zéro (c'est-à-dire en amenant ses deux poignées au contact).

DISCUSSION

Nos premières expériences avec le système nous ont révélé les points suivants :

Il n'est pas nécessaire d'activer le système avant de travailler sur la surface pour bénéficier de ses services. Les utilisateurs peuvent immédiatement travailler sur la surface sans avoir à interrompre leur discussion pour activer le système. Plus tard, quand le besoin des services électroniques se fait sentir, le système est activé afin de capturer l'information qui a été produite. Un autre exemple est le cas où des personnes sont face à une surface remplie du travail d'autres personnes. Notre système peut être activé pour sauvegarder ce travail avant d'effacer la surface. L'activation "a posteriori" est une qualité importante de notre système qui n'est pas possible avec les systèmes fondés sur la capture de la trajectoire des stylos.

Le mécanisme de sélection est rapide et facile à exécuter. C'est une solution élégante à un problème difficile : le système voit de nombreux jetons identiques dont beaucoup sont en mouvement. Il est difficile dans ces conditions d'identifier ceux qui doivent intervenir dans une sélection. Nous devons imposer à l'utilisateur d'effectuer un geste explicite pour identifier le début d'une sélection. Nous avons choisi la métaphore du "chewing gum" : le système réagit comme s'il y avait de la gomme sur les jetons. En mettant les jetons en contact on peut étirer la gomme (qui représente dans notre cas la boîte de sélection). Les geste est intuitif est peut s'exécuter rapidement. De plus, ce

mécanisme autorise le changement de taille et le déplacement de la sélection en parallèle du fait de l'interaction à deux mains.

Nous avons également observés certains problèmes d'utilisation. La caméra bouge et effectue plusieurs plan serré sur la surface pour capturer les inscriptions en haute définition. La procédure peut durer jusqu'à 1 minute si la totalité de la surface doit être capturée. Pendant ce temps le suivi de jeton n'est plus actif car la caméra n'a plus la totalité de l'espace de travail dans son champ de vision. De fait, les utilisateurs ne peuvent plus interagir avec le système. Il est clairement nécessaire de mettre en oeuvre deux caméra : l'une dédiée au suivi des jetons et l'autre à la capture. En fait, l'utilisation d'un appareil photo numérique haute résolution remplacerait avantageusement la caméra

de capture puisqu'elle fournirait la haute résolution en une seule image.

RÉFÉRENCES

1. The Mimio, *Virtual Ink Corporation*,
<http://www.mimio.com/>
2. Mynatt, E. D. Igarashi, T. Edwards, W. K. and LaMarca, A. Flatland: New Dimensions in Office Whiteboards, in *Proceedings of CHI '99* (Pittsburgh, PA, May 1999), ACM Press, 346-353.
3. Saund, E. Image Mosaicing and a Diagrammatic User Interface for an Office Whiteboard Scanner. *Web site*, Xerox Palo Alto Research
<http://www.parc.xerox.com/spl/members/saund/>
4. The Smartboard, *Smart Technologies*,
<http://www.smarttech.com/>
5. Stafford-Fraser, Q. and Robinson, P. BrightBoard: A Video-Augmented Environment. in *Proceedings of CHI'96* (Vancouver, BC, April 1996), ACM Press, 134-141.