Zeitschrift: Bulletin des Schweizerischen Elektrotechnischen Vereins, des Verbandes Schweizerischer Elektrizitätsunternehmen = Bulletin de l'Association suisse des électriciens, de l'Association des entreprises électriques suisses

Herausgeber: Schweizerischer Elektrotechnischer Verein ; Verband Schweizerischer Elektrizitätsunternehmen

Band: 70 (1979)

Heft: 17

Artikel: Traitement et traduction tactile d'images pour prothèse visuelle

Autor: Pun, T. / Coulon, F. de

DOI: https://doi.org/10.5169/seals-905430

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 24.08.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Traitement et traduction tactile d'images pour prothèse visuelle

Par T. Pun et F. de Coulon

616–089.28:617.75

Un grand nombre de recherches sont menées dans le but de remplacer le sens visuel des aveugles; une approche possible est de substituer à ce sens déficient celui du toucher. L'objectif principal du projet décrit est la recherche de méthodes de traitement numérique d'images permettant de réduire la complexité de la scène observée tout en conservant l'information la plus significative. C'est ainsi une sorte d'esquisse binaire de la scène perçue par un capteur opto-électronique qui sera affichée par un transducteur électromécanique approprié.

Es gibt viele Forschungsprojekte mit dem Ziel, blinden Personen das «Sehen» zu ermöglichen. Ein denkbarer Weg besteht darin, das mangelnde Sehvermögen durch den Tastsinn zu ersetzen. Der Hauptzweck des hier beschriebenen Projektes besteht in der Untersuchung von Methoden der numerischen Bildverarbeitung zur Vereinfachung von Bildern unter Beibehaltung der bedeutsamen Information. So wird eine Art binäre Skizze des von einem opto-elektronischen Messwertgebers beobachteten Bildes erhalten, das mit einem geeigneten elektromechanischen Wandler dargestellt und dem Tastsinn zugänglich gemacht werden kann.

1. Introduction

La traduction sous une forme tactile facilement interprétable de certaines informations visuelles serait une aide inappréciable pour les handicapés de la vue (amélioration des possibilités d'action et de travail individuel, accès à une documentation élargie, etc.). A cet égard l'utilisation conjointe de méthodes avancées de traitement d'images, de capteurs électro-optiques et de processeurs électroniques miniaturisés pourrait conduire au développement d'une prothèse visuelle «intelligente», capable de sélectionner l'information la plus significative et de faciliter ainsi la tâche d'interprétation.

Une recherche dans cette direction a été lancée récemment au Laboratoire de traitement des signaux (LTS) de l'EPF-Lausanne avec l'appui du Fonds National Suisse de la Recherche Scientifique. Elle a été motivée par les travaux d'une Commission mixte (Faculté de médecine de l'Université de Lausanne – Centre hospitalier universitaire vaudois – EPFL) de Génie médical qui a, entre autre, conclu à l'intérêt de développer des prothèses sensorielles de substitution se basant sur une restitution tactile de l'information.

Par ailleurs, le traitement numérique des images est un domaine de recherche abordé depuis plusieurs années par le LTS. La simplification de scène par extraction de contours d'objets [2] a déjà été expérimentée avec succès dans le cadre du développement d'un prototype de système de perception visuelle pour robots industriels utilisant une caméra de télévision et une électronique spécialisée [3]. La possibilité d'exploiter ce type de traitement d'images dans le contexte de la traduction élémentaire du sens visuel en sens tactile a initialement fait l'objet d'un projet d'étudiant [1]. Le développement de traducteurs tactiles d'information fait actuellement l'objet d'études au Laboratoire de calculatrices digitales et à l'Institut de microtechnique de l'EPFL.

L'objectif principal de la recherche en cours est la mise au point de méthodes de traitement numérique d'images permettant de réduire la complexité de la scène observée tout en conservant l'information la plus significative. C'est ainsi une sorte d'esquisse binaire de la scène qui sera traduite sous une forme tactile mécanique ou électromécanique appropriée.

L'une des difficultés de l'interprétation tactile réside dans la quantité d'information très élevée contenue dans une image normale. Par application de certaines méthodes de traitement numérique de l'information-image (par ex.: extraction et poursuite de contours, filtrage de détails non significatifs, rehaussement du contraste, «squelettisation», etc.), on peut espérer réduire la quantité d'information et, par conséquent, faciliter le travail d'interprétation. Un tel pré-traitement de l'information visuelle fournie par un capteur d'images électro-optique (par ex.: caméra TV), est considéré comme une étape indispensable à la mise au point d'une véritable prothèse visuelle de substitution.

Un tel sens visuel de remplacement serait donc basé sur un traitement de l'information en deux temps:

a) Pré-traitement électronique de l'information visuelle ayant pour principal objet la simplification de l'image par une extraction et une sélection des éléments les plus significatifs du point de vue cognitif de l'aveugle et leur affichage sous une forme tactile appropriée (images en relief).

b) Exploration tactile (par ex.: digitale) et interprétation faisant appel à l'intelligence et à la mémoire de l'individu, c'est-à-dire basée sur une procédure d'apprentissage préalable.

La possibilité d'une intervention directe du mal-voyant sur le type de pré-traitement effectué ouvre également la voie à des possibilités d'interprétations interactives.

Les méthodes utilisées seront comparées d'une part en fonction de tests d'interprétation effectués par des aveugles, grâce à l'appui de la Clinique ophtalmologique universitaire de Lausanne et, d'autre part, sur la base de leur complexité et de leur aptitude à être mises en œuvre à l'aide d'une électronique miniaturisée pouvant conduire au développement d'une éventuelle prothèse de substitution.

2. Principe d'un système d'aide aux aveugles

On peut caractériser un système d'aide aux aveugles par le schéma fonctionnel simplifié de la figure 1. On distingue ainsi les trois principales étapes séparant l'acquisition de l'information visuelle de sa restitution:

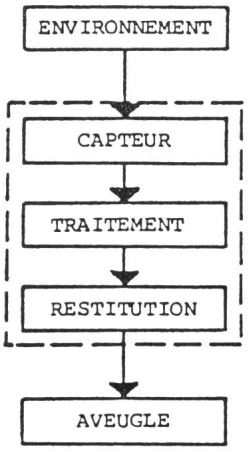

Fig. 1
Schéma bloc d'un système d'aide aux aveugles

– le *capteur,* qui peut être optique (caméra), à rayonnement (faisceau directif de type sonar), tactile (canne blanche), voire même inexistant dans le cas, par ex., d'une machine à écrire parlante;
– le *système de traitement* de l'information provenant du capteur qui peut être matériel (électronique, microélectronique) ou humain (l'handicapé interprète lui-même ce qu'il perçoit);
– Le *système de restitution,* par ex. tactile, auditif, etc.

Selon le type de prothèse, il peut y avoir disparition d'une ou plusieurs de ces étapes.

Dans ce qui suit, on passe d'abord en revue les diverses aides pour aveugles existantes ou en cours de développement. Puis suit une présentation de l'état des recherches effectuées au LTS. On termine en mentionnant les types de transducteurs tactiles dont l'utilisation semble possible.

3. Aides pour aveugles existantes ou en cours de développement

Il en existe actuellement un assez grand nombre, dont malheureusement peu sont vraiment efficaces [4; 5]. Pour les décrire, on les classe selon le mode de restitution à l'aveugle, c'est-à-dire selon la façon dont l'information est retransmise après son traitement. Des cinq sens perceptifs de l'homme (vue, odorat, goût, ouie, toucher) seuls les deux derniers peuvent servir de canal à cette information. On peut encore utiliser directement le système nerveux central, en le stimulant de manière adéquate.

3.1 Restitution auditive

Il faut générer un ou plusieurs sons qui sont une représentation de l'image observée. Dans des cas particuliers où celle-ci est simple à décrire, il est actuellement possible de générer directement des mots; ceci est, par ex., appliqué dans la reconnaissance de caractères.

Dans les *systèmes à faisceau électromagnétique*, une onde électromagnétique directive est émise, par ex. des ultrasons ou de la lumière cohérente. La présence d'obstacles provoque une réflexion de l'onde incidente. Celle-ci est captée et traduite en un signal acoustique dont l'intensité peut varier avec la proximité de l'obstacle, permettant ainsi de localiser l'objet; de tels dispositifs existent sur des cannes et sur des lunettes spéciales. Ils ne fournissent aucune information sur la taille ou la nature de l'obstacle, et leur directivité empêche d'avoir une perception d'ensemble de l'environnement.

Pour donner une information plus globale à l'handicapé, on peut essayer d'analyser l'ensemble de l'image de façon à générer deux signaux acoustiques indépendants soit un *signal stéréophonique*, qui dépend du nombre et de la disposition des points dans la scène observée, et peut, par ex., être donné par une analyse des lignes de celle-ci. Les travaux effectués dans ce domaine [6] avec des images simples telles des cercles et des polygones montrent que le grand (et insurmontable?) problème de cette approche est la difficulté d'interprétation des signaux.

La *synthèse de parole* a été développée pour la reconnaissance de textes, par caractères individuels ou éventuellement par mots entiers. Il existe actuellement des machines à écrire parlantes [7], ayant même certaines possibilités de répétition facilitant l'édition de textes. On trouve également sur le marché des machines à calculer parlantes, permettant ainsi aux aveugles de les utiliser. Les recherches dans ce domaine portent surtout sur la synthèse de phrases complètes. Celle-ci s'effectue par génération de phonèmes qui ont été au préalable répertoriés et mémorisés; il faut également un circuit de simulation du canal vocal, afin de donner à la voix électronique les caractéristiques d'une voix humaine. Ces systèmes sont malheureusement limités aux applications où n'apparaît que du texte.

3.2 Stimulation directe du système nerveux

Il s'agit d'implanter dans les centres de la vision du cortex cérébral des électrodes qui génèrent des sensations lumineuses sur la rétine par envoi d'impulsions électriques. Il faut noter que cette technique n'est possible que si le nerf optique n'est pas lésé.

Les principales recherches dans ce domaine sont effectuées à l'Université de Salt Lake City (Utah, USA). Il semble qu'il soit possible de lire du Braille (transmis par les phosphènes) à une vitesse de 30 lettres par minutes [8]. Il faut cependant remarquer qu'un caractère braille nécessite considérablement moins de points (2×3) que toute une image. Les problèmes techniques sont extrêmement nombreux et complexes. Les interventions directes sur les centres de commande du cerveau posent par ailleurs des problèmes d'éthique médicale.

3.3 Restitution tactile

Les aides *sans traitement de l'information* consistent à imprimer en relief sur un support adéquat une information interprétable par «lecture» tactile avec les extrémités des doigts. Les plus connues de ces méthodes sont les retranscriptions de textes, soit en codant les caractères en Braille, soit tout simplement en les imprimant en relief. Il existe également des plans de villes ou de magasins en relief, sur lesquels des signes spéciaux symbolisent les passages pour piétons, les escaliers, les entrées d'immeubles etc. [9]. Citons aussi la canne blanche, qui en plus de prévenir les gens de l'infirmité dont souffre son porteur, permet la détection d'obstacles et informe de la nature du terrain. Dans tous ces exemples, il s'agit de systèmes où l'handicapé interprète lui-même l'information.

Un traitement élémentaire est fourni par des appareils du type de l'Optacon: c'est un instrument de lecture de la taille d'un enregistreur portable, comportant une tête de lecture mobile. On promène celle-ci sur une ligne du texte, lettre par lettre. La zone observée est analysée par une matrice de 6×24 cellules photosensibles. De façon correspondante, l'aveugle a son index sur un rectangle de 6×24 petites pointes. Si l'un des points de l'image est sombre, la pointe correspondante vibre sous le doigt. Des calculatrices tactiles, ainsi que des systèmes de lectures de textes (les livres sont au préalable enregistrés en Braille sur bande magnétique) sont également équipés de ce type de transducteur à pointes vibrantes.

Des tentatives ont été faites pour présenter l'ensemble d'une image sur une grande matrice de pointes vibrantes. Les travaux dans ce domaine, tels ceux de *C. C. Collins* [10] et de *P. Bachy-y-Rita* [11] mettent cependant en évidence la difficulté de construire une matrice de 100×100 pointes vibrantes. Un certain nombre de solutions possibles pour réaliser un tel transducteur sont présentées au chapitre 5. Malgré les difficultés de réalisation, ce type de transducteur tactile présente par rapport aux transducteurs auditifs ou implantés deux avantages: sa simplicité de mise en œuvre et sa plus grande facilité d'apprentissage pour l'utilisateur. Pour ces raisons, c'est ce type d'interface homme-machine qui semble souhaitable.

Une méthode plus poussée est celle de la *transformation de l'image*, où celle-ci est traitée de manière à faire ressortir les caractéristiques principales des objets qui la composent, puis est transmise à l'aveugle par l'intermédiaire d'une matrice tactile. C'est cette approche qui est abordée dans les travaux du LTS.

4. Travaux en cours à l'EPFL

L'étude actuelle est essentiellement consacrée aux méthodes de traitement par ordinateur de l'image, sans se préoccuper de son acquisition (on travaille à partir de photos). Pour l'instant, le mode de restitution tactile envisagé (fig. 1) est du type binaire, ce qui implique que le résultat du traitement soit une image à deux niveaux: blanc ou noir. L'image est *numérisée*, c'est-à-dire représentée par un nombre fini de points pouvant chacun prendre un nombre fini de valeurs de luminance, appelées *niveaux de gris*.

Dans ce qui suit, la photo originale (fig. 2) a été numérisée en 128×128 points, chacun pouvant prendre 256 valeurs de gris entre le niveau 0 (blanc) et le $255^{\text{ème}}$ (noir). On peut noter sur cette photo certains défauts typiques d'une image courante: reflets (sur le pot à lait et le couteau), ombres etc.

4.1 L'image recherchée

Des tests ont été effectués avec la collaboration d'enfants de l'Institut pour jeunes aveugles de Lausanne afin de déterminer quel type d'image est le plus aisé à interpréter tactilement: avec beaucoup de détails, ou au contraire très dépouillée? Il est apparu sans conteste que les objets doivent être le plus *stylisés* possible (fig. 3). Tout ce qui, pour un voyant, apporte des informations supplémentaires (reflets, ombres, texture ...) ne constitue pour un handicapé de la vue que du *bruit*, c'est-à-dire des perturbations gênant l'acquisition de l'information. Ces tests ont soulevé un autre problème: chez les enfants aveugles de naissance, la notion de *relief* est inconnue. Lorsqu'on dessine un verre dont l'ouverture forme une ellipse, l'enfant ne comprend pas ce qu'elle représente. Dans une telle situation, où il semble impossible d'effectuer un traitement automatique qui remplacerait par ex. le verre de profil, l'apprentissage de l'utilisateur devra se substituer à sa construction mentale du monde extérieur. Deux aspects peuvent cependant faciliter l'interprétation d'une image présentée tactilement:

– l'aide du *contexte*, qui fait que quelqu'un se sachant dans sa salle de bain ne s'attend pas à y trouver une brouette.

– la possibilité d'avoir successivement *plusieurs représentations* d'une même scène sous des angles différents.

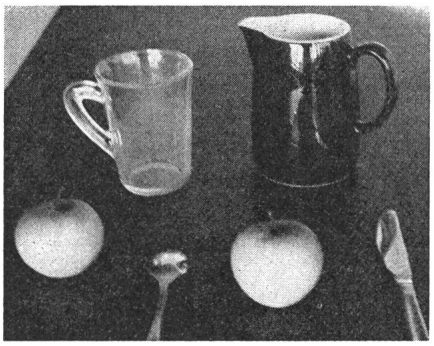

Fig. 2 Une des photos utilisées pour ce travail

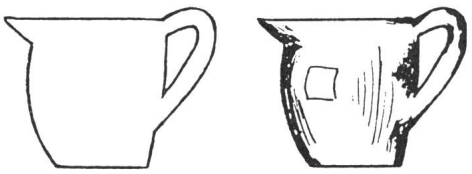

Fig. 3 Différences entre un pot représenté de façon stylisée ou avec beaucoup de détails

4.2 Etapes du traitement

Rappelons que l'on cherche à obtenir une image à deux niveaux, stylisée, avec le moins de bruit possible et tout ceci avec la contrainte d'un traitement numérique le plus rapide possible. Les principales étapes du traitement, détaillées dans ce qui suit, sont: une extraction de contours, faisant ressortir les zones de transitions des niveaux de gris, un seuil donnant une image à deux niveaux, une «poursuite» éliminant du bruit et «squelettisant» les contours, un rehaussement de l'image afin de la rendre plus interprétable.

4.3 Extraction de contours [2]

Filtrage passe-haut: De la même façon que l'on définit une transformée de Fourier pour un signal dependant du temps, un signal bidimensionnel (une image) peut être caractérisé par son spectre. On parle alors de fréquences *spatiales*, exprimées en m^{-1}. Un signal ayant de nombreuses transitions de niveaux de gris est caractérisé par un spectre à hautes fréquences spatiales. Le contour d'un objet, sur une image en noir et blanc, est une zone de transition des niveaux de gris. On va donc la faire ressortir par un filtrage passe-haut.

Il existe beaucoup d'opérateurs effectuant ce filtrage. On peut les classer en trois catégories principales: les opérateurs de type «gradient» donnés par une dérivée locale première (discrète) de la fonction «niveaux de gris», les opérateurs de type Laplacien donnés par une dérivée seconde et les opérateurs heuristiques, découverts et appliqués lors de recherches sur un type d'image bien déterminé. Pour des raisons de simplicité, donc de rapidité de calcul on s'intéresse particulièrement à trois opérateurs de type gradient.

Opérateur de Roberts [2]: On attribue au point x_R (fig. 4a) une valeur donnée par une combinaison des valeurs de points de l'image – multipliés par des coefficients appropriés – pris sur une fenêtre d'observation de 2×2 points:

$$x_R = |a| + |b| \tag{1}$$

où a et b sont définis à la figure 4a. Il s'agit d'une expression simplifiée de l'opérateur gradient pour lequel:

$$x_G = \sqrt{a^2 + b^2} \tag{2}$$

On calcule ainsi la valeur d'un certain point, on décale la fenêtre d'une colonne et on recommence, ceci pour chaque ligne (convolution bi-dimensionnelle). Appliqué à l'image de la figure 2, l'opérateur de *Roberts* donne l'image de la figure 5a.

Opérateur de Mero-Vassy: La fenêtre de calcul est plus compliquée (fig. 4b); elle a pour but de permettre l'extraction de dérivées plus directionnelles, donc de mieux tenir compte des transitions de niveaux dans toutes les directions. On attribue à x_{MV} la valeur

$$x_{MV} = |a| + |b| \tag{3}$$

où a et b sont définis à la figure 4b. L'image obtenue est représentée à la figure 5b. Elle ne diffère guère de celle obtenue avec l'opérateur de *Roberts*.

Opérateur de Sobel: La fenêtre de calcul porte ici sur 9 points. On a à nouveau

$$x_S = |a| + |b| \tag{4}$$

où a et b sont définis à la figure 4c, ce qui donne l'image de la figure 5c. On constate un épaississement des contours par rapport aux deux images précédentes. Ceci est dû au fait que les coefficients de pondération (1; 2; 1) de l'opérateur constituent une fenêtre passe-bas afin de diminuer l'influence du bruit.

L'opérateur d'extraction de contours donne des images dans lesquelles ressortent mieux les objets, mais qui sont également très «bruitées». La méthode suivante permet d'éliminer partiellement ces perturbations.

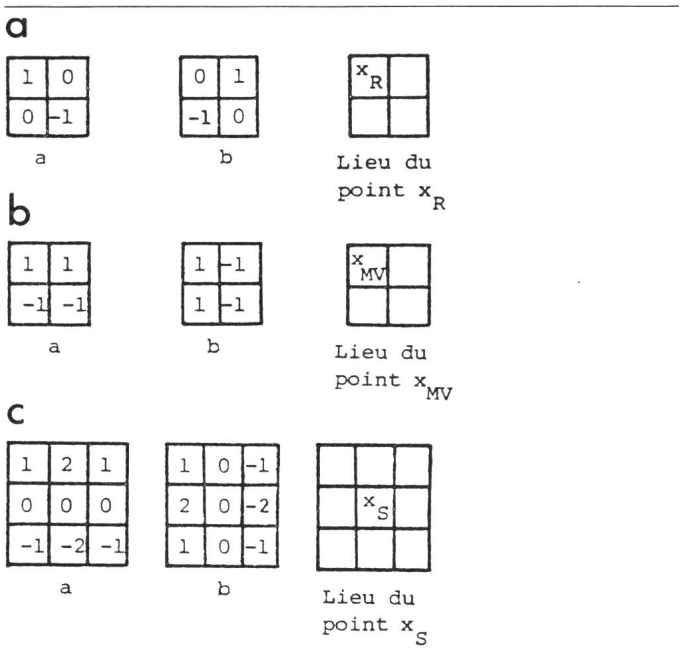

Fig. 4 Fenêtres de filtrage passe-haut, et localisation dans la fenêtre du résultat du filtrage pour les opérateurs de Roberts (a), Mero-Vassy (b) et Sobel (c)

4.4 Seuils

Après le filtrage passe-haut, on a toujours une image à plusieurs niveaux. On désire la transformer en une image à deux niveaux (blanc ou noir), et avec des contours «squelettisés». L'opérateur de décision logique transformant une image à plusieurs niveaux en image avec des points soit blancs, soit noirs, est appelé *seuil*. Si $z(x, y)$ est le niveau de gris du point (x, y), on obtient après application du seuil s l'image caractérisée par les niveaux $z_S(x, y)$ valant

$$z_S(x, y) = \begin{cases} 1 \text{ si } z(x, y) > s \\ 0 \text{ si } z(x, y) \leq s \end{cases} \tag{5}$$

On s'intéresse ici à la détermination automatique du seuil en fonction des niveaux de gris de la zone où il doit être appliqué. Cette opération de seuil et de «squelettisation» peut être effectuée de deux manières: d'une part au cours d'une opération de «poursuite des extrema locaux» sur l'image multiniveaux, en cherchant à former un «segment» de points connexes ayant une amplitude supérieure au seuil de la zône où s'effectue cette poursuite; d'autre part en appliquant d'abord le seuil, ce qui donne une image à deux niveaux, puis en faisant une «poursuite des contours» sur celle-ci. A cause de la simplicité qu'elle présente, la seconde solution est préférable.

Dans ce qui suit, on présente les résultats obtenus avec deux seuils, appliqués sur l'image filtrée passe-haut par l'opérateur de *Sobel* (fig. 5c).

La principale information dont on dispose pour caractériser une zone de l'image est son *histogramme des niveaux de gris*, qui est la densité de probabilité de la variable aléatoire discrète $z(x, y)$. La figure 6 présente un tel histogramme; on distingue des extrema locaux correspondant à des niveaux de gris plus fortement représentés que les autres sur l'image considérée.

Le *seuil «moyenne pondérée»* est déterminé par une moyenne pondérée des niveaux de gris correspondant aux extrema

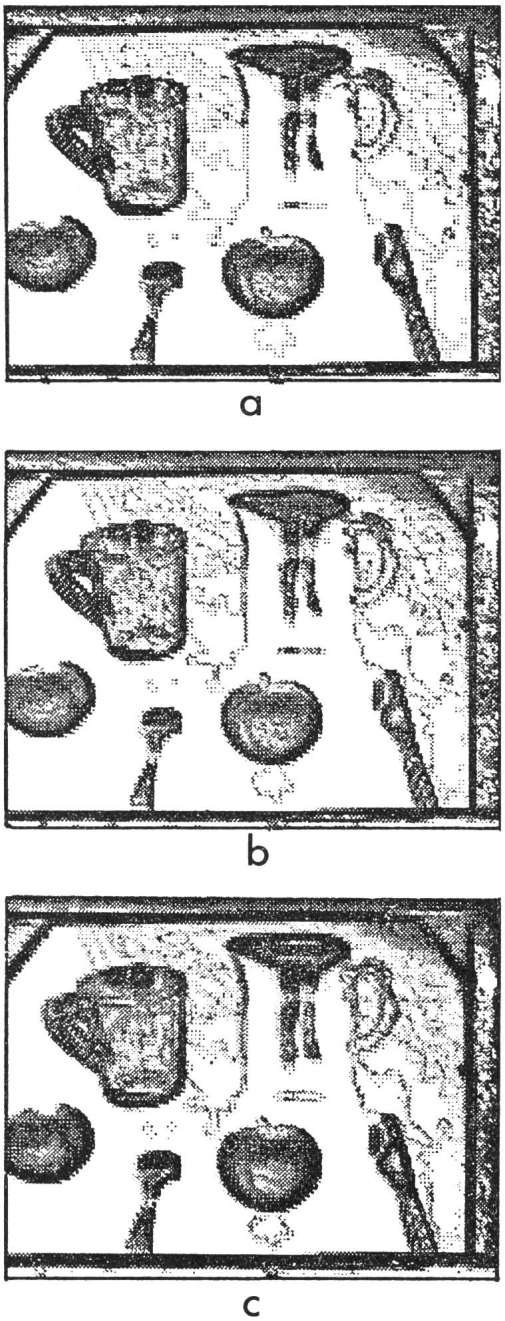

Fig. 5 Images obtenues après filtrage passe-haut de l'image originale par les opérateurs de Roberts (a), Mero-Vassy (b) et Sobel (c)

locaux. Le facteur de pondération dépend de la densité du niveau considéré. Soient z_i le niveau de gris de l'extremum i, et A_i la densité correspondante, A_{max} la plus grande densité de l'histogramme et M le nombre d'extrema locaux. Alors

$$s_{MP} = \frac{1}{M} \sum_{i=1}^{M} z_i \log \left(1 + 9 \cdot \frac{A_i}{A_{max}}\right) \tag{6}$$

Si $M = 1$, $A_i = A_{max}$ et $s_{MP} = z_i$. Si $A_1 = A_2 = ... = A_{max}$,

$s_{MP} = \frac{1}{M} \sum_{i=1}^{M} z_i$ et on a une moyenne arithmétique.

Un paramètre important pour le seuil est la *taille* de la zone sur laquelle est calculé et appliqué le seuil. En effet, si on détermine l'histogramme sur une petite zone, on a peu d'extrema locaux. Ceux correspondant au bruit ont donc une influence plus grande que dans l'histogramme d'une grande zone où les extrema locaux sont nombreux. On peut dire que dans une petite zone, le bruit influence autant le seuil que l'information utile, alors que dans une grande zone, les contours donnent (après filtrage passe-haut) des extrema situés surtout dans les hauts niveaux, et ont ainsi une importance prépondérante. C'est effectivement ce que l'on constate sur les images 7a, b et c.

On voit également que ce seuil «moyenne pondérée» laisse subsister beaucoup de bruit, et n'est donc pas très efficace.

Seuil «entropie»: En théorie de l'information, on définit l'entropie d'une source, ou «quantité d'information moyenne» par symbole, par:

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \quad \text{shannon/symbole} \tag{7}$$

où p_i est la probabilité d'apparition d'un des n symboles de la source. La quantité d'information moyenne d'un point d'une de l'image dépend donc directement des valeurs de l'histogramme des niveaux de gris:

$$H = -\sum_{i=1}^{n} \left(\frac{A_i}{T}\right) \log_2 \left(\frac{A_i}{T}\right) \quad \text{shannon/symbole} \tag{8}$$

où A_i est la densité d'un des n niveaux de l'histogramme et T le nombre total de points de la zone.

Le principe du seuil est de maximiser la quantité d'information H' de l'image *après* son application, donc sur l'image à deux niveaux. Pour cela, on place le seuil au niveau s tel que

$$H_1 = -\sum_{i=1}^{s} \left(\frac{A_i}{T}\right) \log_2 \left(\frac{A_i}{T}\right) = aH \tag{9}$$

$$H_2 = (1 - a)\, H = H - H_1 \tag{10}$$

On subdivise ainsi l'histogramme, considéré comme une source à n symboles, en deux sous-ensembles contribuant à l'entropie totale dans les proportions a et $(1 - a)$. Il faut donc déterminer la valeur de a maximisant H' (entropie de la zone après application du seuil). Soient A_0' et A_1' les densités des niveaux 0 et 1 sur l'image à deux niveaux.

$$H' = -\left(\frac{A_0'}{T}\right) \log_2 \left(\frac{A_0'}{T}\right) - \left(\frac{A_1'}{T}\right) \log_2 \left(\frac{A_1'}{T}\right) \tag{11}$$

H' est maximum quand $A_0' = A_1'$,

$$\text{d'où} \sum_{i=1}^{s} A_i = \sum_{i=s+1}^{n} A_i \tag{12}$$

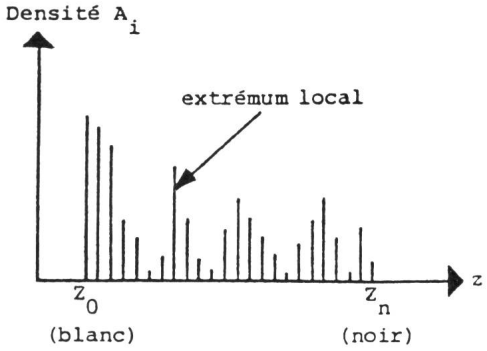

Fig. 6 Histogramme des niveaux de gris d'une image comportant n extréma locaux

A_i densité d'un des n niveaux de gris Z_i
Z_0 et Z_n niveaux extrêmes (appelés par convention «blanc» et «noir»)

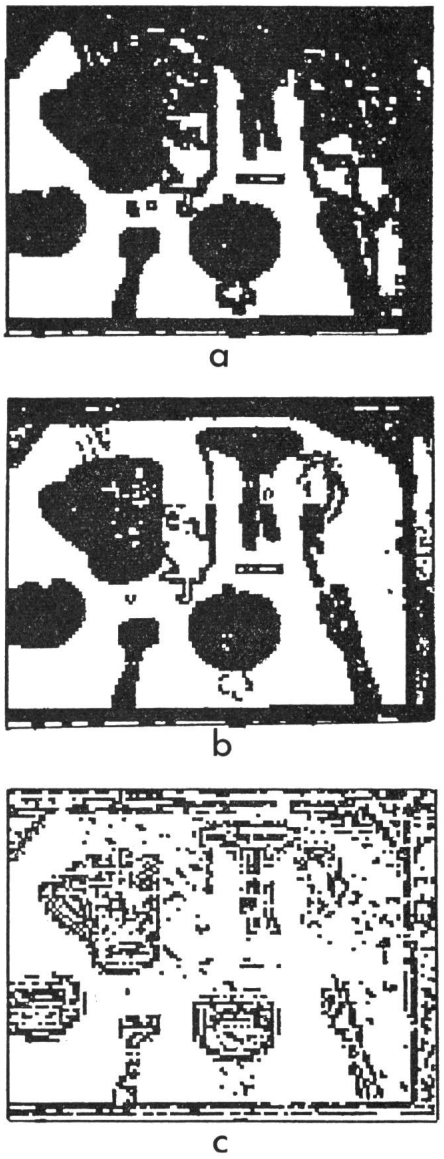

Fig. 7 Images obtenues après filtrage passe-haut par l'opérateur de Sobel, et application du seuil «moyenne-pondérée» calculé avec un histogramme déterminé sur toute l'image (a) $^1/_{16}$ème de l'image (b), $^1/_{4096}$ème de l'image (c)

D'autre part:

$$\left.\begin{array}{l} H_1 = aH \\ H_2 = (1-a)\,H \end{array}\right\} \Rightarrow H_1 = bH_2 \text{ avec } b = \frac{a}{1-a} \tag{13}$$

Avec (9) et (10)

$$\sum_{i=0}^{s} \left(\frac{A_i}{T}\right) \log_2 \left(\frac{A_i}{T}\right) = b \sum_{s+1}^{n} \left(\frac{A_i}{T}\right) \log_2 \left(\frac{A_i}{T}\right) \tag{14}$$

d'où

$$\sum_{i=0}^{s} A_i \log_2 A_i = b \sum_{s+1}^{n} A_i \log_2 A_i + (\log_2 T) \cdot \left(\sum_{0}^{s} A_i - b \sum_{s+1}^{n} A_i\right) \tag{15}$$

En tenant compte de (12), la seule solution n'imposant pas de condition sur les A_i est $b = 1$, ce qui implique $H_1 = H_2 = H/2$, équivalent à (12).

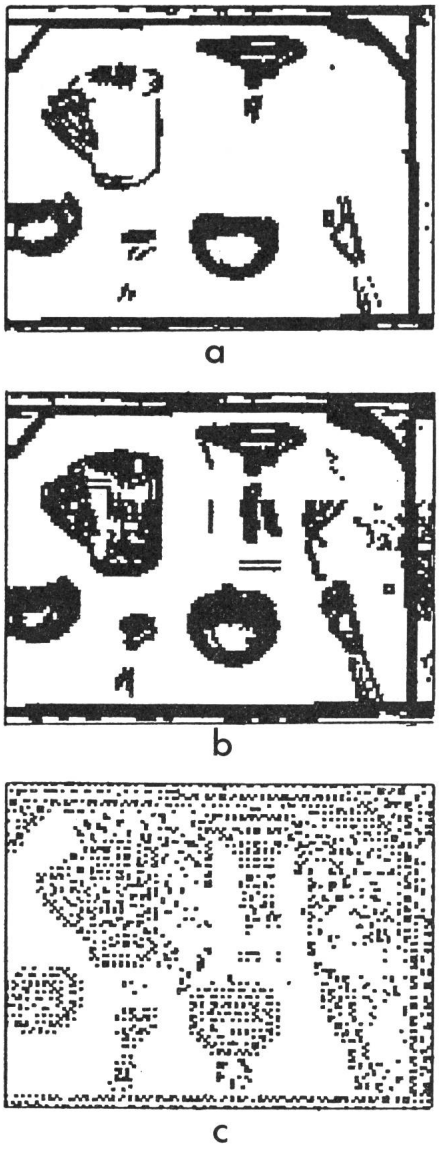

Fig. 8 **Images obtenues après filtrage passe-haut par l'opérateur de Sobel et application du seuil «entropie» calculé avec un histogramme déterminé sur toute l'image (a), $^1/_{16}$ème de l'image (b), $^1/_{4096}$ème de l'image (c)**

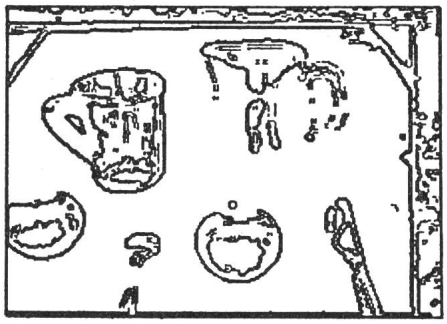

Fig. 9 **Image obtenue après double filtrage passe-haut par l'opérateur de Sobel et seuil «entropie» (histogramme de $^1/_{16}$ème d'image)**

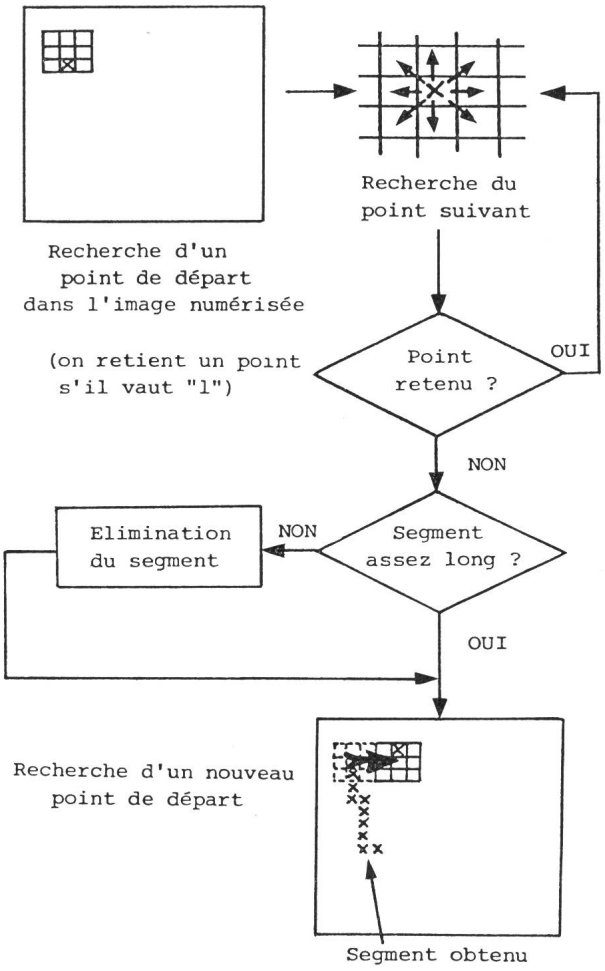

Fig. 10 **Organigramme de la poursuite des contours (formés de points à «1») avec élimination des segments trop courts**

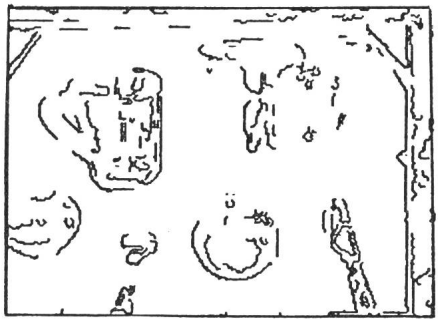

Fig. 11 **Image de la figure 9 sur laquelle une poursuite des contours avec élimination des segments de moins de 5 points à été pratiquée**

Pour maximiser l'entropie de l'image à deux niveaux, il faut que le seuil s soit tel que (12) soit vérifiée. On associe ainsi la *même quantité d'information* aux points noirs et aux points blancs.

Les figures 8a, b, c présentent l'effet de ce seuil en déterminant l'histogramme sur toute l'image, $^1/_{16}$ et $^1/_{4096^{\text{ème}}}$ de celle-ci, respectivement.

Double extraction de contours: L'apparition de grandes plages noires suggère d'appliquer à nouveau un filtrage passe-haut sur les images précédentes. Le résultat de cette opération sur la figure 8a, est présenté à la figure 9. On constate bien que la majeure partie du bruit a disparu. On a aussi perdu de l'information, ce qui montre les dangers d'une telle opération. Il y a là un compromis à réaliser entre cette perte d'information et la diminution du bruit.

Il faut cependant noter que la perte d'information n'a pas les mêmes conséquences que dans le cas, par ex., de codages d'image pour leur transmission où il faut pouvoir reconstituer tout ce que l'on avait au départ. Dans le cas d'une prothèse, il y a toujours la faculté d'interprétation de l'utilisateur pour permettre d'améliorer le résultat global.

4.5 Poursuites des contours

Le but de cette poursuite est d'éliminer les segments trop courts, que l'on suppose être du bruit, et de «squelettiser» les autres segments.

L'algorithme de cette opération est présenté à la figure 10. Sa simplicité permet une programmation plus efficace que si la poursuite est effectuée sur l'image à plusieurs niveaux de gris. La figure 11 montre le résultat d'une telle poursuite au cours de laquelle on élimine les segments de moins de 5 points de longueur. Un autre critère possible d'élimination du bruit est de supprimer les segments ayant de trop grande variations de courbure. Mais dans le cas d'une image numérisé, les critères de courbure sont assez difficiles à déterminer; les résultats obtenus pour l'instant avec ce type de conditions ne sont pas meilleurs qu'avec la poursuite simple.

4.6 Rehaussement

En éliminant beaucoup de bruit par les algorithmes précédents, on a aussi éliminé de l'information utile, par ex. en fragmentant les contours des objets. Le rehaussement vise à pallier à ceci afin de rendre l'image plus interprétable. Une approche est de remplacer les segments par des arcs de coniques, et de les prolonger de part et d'autre des interruptions de contours. Mais les recherches effectuées dans cette voie n'ont pas encore donné de résultats très probants.

5. Possibilités de réalisations de matrices tactiles

D'après [10], le champ de vision de l'œil a une définition d'environ 200 × 200 points. Pour des raisons de temps de calcul ainsi que de place mémoire nécessaire, nous nous contentons actuellement d'images de l'ordre de 100 × 100 points. Même dans ce cas, les problèmes technologiques de réalisation d'une matrice tactile de 10000 éléments sont considérables. Une façon de tourner ce problème serait de ne présenter à l'aveugle qu'une partie de l'image traitée, qu'il pourrait choisir en actionnant un «manche à balais».

Les différents types de transducteurs tactiles rencontrés dans la littérature sont différenciés par le mode de transmission du niveau «noir». Citons:

- matrices à jets d'air [12]: mise en œuvre extrêmement lourde, faible définition spatiale
- matrices de contacts électriques [13]
- matrices de pointes vibrantes, comme pour l'Optacon
- combinaison de vibrations mécaniques et de stimulation électrique [14]
- matrices de métaux dits «à mémoire» (effet de dilatation thermique): gros problèmes de temps de relaxation
- matrices exploitant l'effet Johnson-Rabek, qui est l'attirance électrostatique dans une jonction semi-conducteur métal, la peau jouant le rôle du semi-conducteur.

Les deux derniers types de matrice font actuellement l'objet de recherches à l'Institut de microtechnique de l'EPFL. Parmi toutes ces possibilités aucune n'a encore donné des résultats très positifs.

Bibliographie

[1] *T. Pun:* Traitement numérique d'images pour prothèse visuelle. Lausanne, Projet de diplôme, hiver 1978. Lausanne, Ecole Polytechnique Fédérale de Lausanne, Laboratoire de Traitement des Signaux, 1978.

[2] *P. Kammenos:* Extraction de contours en traitement électronique des images. I: Principaux opérateurs de traitement. Bull. ASE/UCS 70(1979)11, p. 525...531.

[3] *J.-M. Zürcher:* Extraction de contours en traitement électronique des images. II: Processeur spécialisé pour signal vidéo. Bull. ASE/UCS 70(1979)11. p. 532...536.

[4] Prosthetics aids for the blind. Symposium of the 78th Annual Meeting of the American Academy of Ophthalmology and Otolaryngology, Dallas, 16...20 sept. 1973. Transactions of the American Academy of Ophthalmology and Otolaryngology 78(1974)5.

[5] *J.F. Cuendet:* Moyens auxiliaires modernes d'aide aux handicapées visuels. Medecine et Hygiène –(1977)1234.

[6] *R.M. Fish:* An audio display for the blind. Trans. IEEE BME 23(1976)2, p. 144...154.

[7] *C.Y. Suen* and *M.P. Beddoes:* Development of a digital spelled speech reading machine for the blind. IEEE Trans. BME 20(1973)6, p. 452...459.

[8] *W.H. Dobelle* a.o.: Braille reading by a blind volunteer by visual cortex stimulation. Neuroprostheses program. UT.84115. Salt Lake City, University Utah, Institute of Biomedical Engineering.

[9] *J.D. Armstrong:* The development of tactual maps for the visually handicapped. Active touch, the mechanism of recognition of objects by manipulation: a multidisciplinary approach. Oxford, Pergamon Press, 1978.

[10] *C.C. Collins:* Tactile television-mechanical and electrical image projection. IEEE Trans. MMS 11(1970)1, p. 65...71.

[11] *P. Bach-y'Rita:* Visual information through the skin. Prosthetics aids for the blind. Symposium of the 78th Annual Meeting of the American Academy of Ophthalmology and Otolaryngology, Dallas 16...20 sept. 1973. Transactions of the American Academy of Ophthalmology and Otolaryngology 78(1974)5, p. 729...740.

[12] *C. Rodgers:* Choice of stimulator frequency for tactile arrays. IEEE Trans. MMS 11(1970)1, p. 5...11.

[13] *M. Solomonow* a.o.: Electrotactile two-point discrimination as a function of frequency, body site, laterality, and stimulation codes. Annals of BME 5(1977)1, p. 47...60.

[14] *H. Ide:* Vibrotactile stimulation by using both the electrical and mechanical stimuli. Med. Biol. Eng. –(1976)7, p. 451...454.

Adresse des auteurs

Thierry Pun, ing. él. diplômé, et Prof. *Frédéric de Coulon,* Laboratoire de Traitement des Signaux, EPFL, Chemin de Bellerive 16, 1007 Lausanne.