

L'INTELLIGENCE ARTIFICIELLE :

L'ORDINATEUR RECONNAIT LES FORMES

(Suite voir n° 1410)

PENDANT longtemps, on a dû se contenter, pour introduire dans les calculatrices les informations qu'on voulait leur faire traiter, des cartes perforées. Cette technique n'était pas sans présenter des inconvénients : il fallait, pour procéder au transfert des informations détenues par le document original, en un document lisible par les machines, recourir à des perforations, ce qui exigeait un personnel spécialisé et un matériel *ad hoc*. Des erreurs pouvaient se glisser dans cette transcription. Enfin, la carte n'étant compréhensible que par la machine, il fallait obligatoirement conserver le document original.

CONSTRUIRE DES MACHINES QUI SAVENT LIRE !

A cette méthode, qui demandait à l'homme de se plier aux exigences de la machine, se sont progressivement substituées d'autres méthodes permettant à la machine de tenir un meilleur compte des habitudes humaines : apprendre à lire par exemple. La machine devait pouvoir saisir l'information déposée sur une feuille de papier, sous la forme de caractères alphanumériques ou encore de graphiques et photographiques. L'opération, logiquement, appelle deux temps : lecture de l'information ; puis identification.

Encre magnétique et codage des chiffres et lettres : l'association de ces deux techniques conduisit à la première génération de machines à lire automatiques.

La seconde génération fut celle de la reconnaissance op-

tique des caractères dactylographiés codés.

L'apparition, au SICOB 66, du lecteur optique IBM 1287, capable de reconnaître toutes combinaisons d'informations numériques, dactylographiées, marquées ou manuscrites, ouvrit l'ère de la troisième génération.

La quatrième génération est en cours de développement : la machine se fait plus ambitieuse, elle reconnaît les informations graphiques.

A chacune de ces générations de machines à lire furent associés des sigles : MICR (Magnetic Ink Character Readers) pour la première ; OMR (Optical Mark Readers) et OCR (Optical Character Recognition) pour les deux générations suivantes ; enfin, IDP (Image Data Processors) pour la dernière génération.

UNE MULTITUDE DE CODES POUR L'OCR

Pour mettre au point et diffuser à une échelle commerciale les lecteurs de la seconde génération, il fallait surmonter de nombreuses difficultés : celle afférentes aux dispositifs de lecture, et surtout celles relatives à la codification. Il fallait définir des caractères qui ressemblent le plus possible aux caractères d'usage courant, et qui se prêtent le mieux à l'identification par lecteur optique.

La logique voulait que les constructeurs s'entendent sur un jeu de caractères pouvant être interprétés par toutes les machines. Ce critère d'universalité s'est heurté à de nombreux obstacles, de sorte qu'aujourd'hui, les codes de base sont aussi

OCR-A

A	B	C	D	E
F	G	H	I	J
K	L	M	N	O
P	Q	R	S	T
U	V	W	X	Y
Z	1	2	3	4
5	6	7	8	9

OCR-B

A	B	C	D	E
F	G	H	I	J
K	L	M	N	O
P	Q	R	S	T
U	V	W	X	Y
Z	1	2	3	4
5	6	7	8	9

Fig. 3

diversifiés que les matériels qui les utilisent.

Les principaux codes ainsi développés sont :

- le NOF (Numeric Optical Font) de la NCR : caractères numériques essentiellement employés dans les caisses enregistreuses, pouvant éventuellement être lus par le périphérique IBM 1287 ;
- les caractères OCR-A, B, C (Fig. 3) développés par l'American National Standards Institute ; ils sont d'un emploi courant, en particulier avec l'UNIVAC 2703, le 3010 de Farrington, le CD 935-1, le lecteur 1287 de IBM, l'OPSCAN 288, le 243 de Honeywell et enfin le Burroughs 9134 ;
- les caractères 7B, 12L et 12F de Farrington : les deux premiers concernent les caractères alphanumériques ; le 12F est purement numérique. Ils sont utilisés sur le matériel Farrington 3010 et peuvent être reconnus par l'IBM 1287. Le lecteur 935-1 de la CDC peut lire les caractères 7B et 12F. Les 7B ont la particularité d'être plus larges et hauts que les autres caractères, ce qui implique l'emploi de machines à écrire spéciales ;
- les caractères 1428, exclusivement numériques ont été conçus par IBM pour son modèle 1428. Ils furent suivis par les caractères 1428E, plus grands que les précédents. Les équipements suivants peuvent, avec le 1428 IBM, reconnaître ces caractères.

tères : Farrington 3010, CDC 935-1, IBM 1287 ;

- les caractères numériques 407 : ils sont exclusivement employés par IBM ;
- les caractères numériques N-2, développés par RCA pour son modèle 70/251 ;
- les caractères numériques H-14, similaires aux 407 de IBM ;
- les caractères E-13B : ils ne sont pas destinés aux lecteurs optiques : ils furent développés avant la mise au point des OCR et ont été employés par les banques, pour le codage magnétique de leurs chèques. Cependant, certains lecteurs optiques reconnaissent ces caractères : d'où leur intérêt pour les banques envisageant l'utilisation de lecteurs optiques.

OCR & TECHNOLOGIE

Un lecteur optique de caractères alphanumériques est composé essentiellement de trois unités fonctionnelles :

- une unité de transport de documents ;
- une unité de balayage ;
- une unité de reconnaissance.

La première unité apporte les documents contenus dans un magasin à l'unité de balayage. Dans la plupart des cas, la méthode de lecture repose sur la différence de réflectivité qu'offrent, sur une surface donnée, les zones qui demeurent blanches et celles qui se trouvent noircies. Explorant cette surface, un fin pinceau lumineux peut être modulé par cette différence, et un analyseur de lumière peut en tirer des déductions.

Trois lecteurs optiques (3010 Farrington ; 1418 et 1428 de IBM) sont pourvus d'une unité de balayage mécanique. Le caractère à lire est fortement éclairé : par l'intermédiaire d'une lentille optique, une image réelle du caractère est projetée sur un disque tournant perforé, derrière lequel se trouve un diaphragme. Ainsi, grâce aux perforations dans le disque, l'image du caractère se trouve échantillonnée ; l'intensité lumineuse recueillie dans chaque échantillon est limitée par le diaphragme. Le récepteur de lumière, derrière le diaphragme est un photomultiplicateur.

La technique de balayage par flying-spot apporte beaucoup plus de souplesse d'emploi que le balayage mécanique. Le spot d'un tube cathodique est dirigé par une lentille optique placée devant l'écran du tube, vers le document à

lire ; la lumière réfléchie est alors captée par un photomultiplicateur, qui délivre un signal électrique proportionnel à l'intensité réfléchie. L'IBM 1287 a un lecteur conçu d'après ces principes.

On peut également envisager l'emploi d'une batterie de cellules photoélectriques pour lire simultanément tous les caractères d'une page entière ; chaque cellule est chargée de détecter la présence d'une petite surface claire ou sombre. Un commutateur électronique, placé derrière les cellules, mesure le courant délivré par chacune des cellules ; ce courant est, bien sûr, fonction de l'intensité lumineuse réfléchie par chaque région élémentaire du document. De nombreux équipements utilisent la méthode de balayage par cellules photoélectriques : UNIVAC 2703, CDC 935-1, OPSCAN 288, Honeywell 243.

Le balayage par tube vidicon représente un moyen entièrement différent de balayage de document : chaque caractère est projeté sur une caméra vidicon de télévision, balayée par un faisceau électrique ; les signaux vidéo résultants sont ensuite quantifiés pour indiquer la présence de surfaces sombres ou claires, localement, sur le document. La RCA a, voici quelques années, développé un tel dispositif.

La comparaison de ces quatre modes de balayage montre l'intérêt présenté par les dispositifs entièrement électroniques. Le balayage mécanique est relativement lent, mais par contre bon marché, et fiable. Le dispositif à flying-spot est classé parmi les unités à moyenne vitesse, avec le tube vidicon. La solution la plus rapide est certainement la rétine à matrices de cellules photoélectriques ; sa vitesse de lecture est de plusieurs milliers de caractères par seconde.

La vitesse de lecture n'est pas le seul paramètre important d'une unité OCR : il ne doit pas y avoir d'erreur de lecture. Or l'une des sources d'erreurs est liée à la nature même de la lumière réfléchie par les surfaces claires ou sombres. L'emploi d'un laser comme source lumineuse devrait améliorer considérablement la fiabilité de la lecture optique ; trois firmes américaines : CDC, Creative Logic Corp. et Optical Business Machines développent de telles machines ; tout récemment, Hitachi a annoncé la commercialisation dès le début 73, d'un lecteur optique à laser, capable de reconnaître 1 800 pages de 2 088 caractères à l'heure (soit environ 1 000 caractères par seconde).

LES UNITES MONOLITHIQUES DE RECONNAISSANCE OPTIQUE

L'optoélectrique devrait conduire à d'importants bouleversements : les fabricants de circuits intégrés s'initient au mariage des transistors MOS intégrés et des matrices intégrées de phototransistors et photodiodes ; le fruit d'une telle association est une unité de lecture optique à balayage interne, sur des cadences extrêmement élevées (ces cadences se chiffrent en mégahertz !), avec une résolution très grande.

Quatre compagnies étudient et commercialisent des matrices de lecteur optique : Reticon, Fairchild et Texas Instruments, aux U.S.A. ; Integrated Photomatrix Ltd, outre-Manche. Les rétines réalisées contiennent entre 64 et 256 éléments photosensibles, et leur coût actuel se chiffre approximativement autour du dollar par élément photosensible.

La possibilité de lire à des cadences très élevées (au moins 10 MHz) pousse les constructeurs à analyser toutes les possibilités offertes par les circuits intégrés : n'oublions pas que le marché des unités OCR devrait inclure les P.T.T., et que les rétines électroniques vont avoir à équiper des machines capables de trier des millions de lettres chaque jour. Aux U.S.A., Reconition Equipment et IBM devraient prochainement livrer à la poste centrale de New York, un tel matériel pour lire les codes postaux et adresses, sur les courriers d'affaires.

Les dispositifs dits « à transfert de charges électriques » constituent une autre voie de développement des rétines électroniques. Aux laboratoires de recherches de Philips, une telle rétine a été construite, associant une rangée de 64 photodiodes à une mémoire « chaîne à seaux », dont la cadence de travail pourrait atteindre 10 mégahertz. Joop Van Santen qui a mis au point cette rétine, en développe une plus importante, sous forme d'une matrice 50 x 50.

Les mémoires « chaîne à seaux » (les anglo-saxons les ont appelées « bucket-brigade » par analogie avec la vieille méthode de lutte contre le feu, dans laquelle des seaux d'eau passaient de main en main), sont en fait des registres à décalage analogique : des charges, stockées dans des condensateurs, sont déplacées vers d'autres condensateurs lorsqu'un signal extérieur au circuit (c'est une horloge) est appliqué sur

un commutateur. Réalisé en circuit intégré, le commutateur est un transistor MOS ou bipolaire, et le condensateur est formé par les capacités entre jonctions et connexions.

Des travaux analogues ont été entrepris aux U.S.A., où les Bell Telephone Laboratories mettent au point des dispositifs « à couplage de charges » (CCD : charge-couples devices) ; pour sa part, la RCA travaille sur des « transistors à charges de surface » (SCT : surface-charge transistor), qui sont considérés comme des CCD modifiés. Enfin, récemment, Fairchild a dévoilé un dispositif CCD amélioré, à « canal enterré » dont la cadence de fonctionnement atteint 20 mégahertz.

Ce dernier résultat est important : les rétines à registres à décalage analogiques peuvent fonctionner à de très grandes vitesses, certainement bien plus élevées que les rétines classiques. Les CCD à canaux enterrés juste sous la surface de la pastille semiconductrice devraient, dans les prochains mois, fonctionner à des cadences voisines de la centaine de mégahertz : on espère même atteindre le gigahertz (soit 1 000 mégahertz) dans les années à venir !

Un dispositif CCD est constitué d'électrodes déposées sur la surface d'une pastille semiconductrice. L'application de tensions électriques adéquates « creusent », au sein du semiconducteur, des puits de potentiel, dans lesquelles peuvent « tomber » des charges, injectées par exemple par des photodiodes. Après la phase « stockage » des charges injectées, prend place la phase « transfert », au cours de laquelle les tensions électriques sont modifiées de telle sorte que les puits de potentiel se trouvent déplacés vers une électrode de sortie ; cette dernière électrode est connectée à une unité de traitement qui compte les paquets de charges contenues dans chaque puits de potentiel.

Dans l'amélioration apportée par Fairchild, les charges se meuvent dans la couche enterrée. Fairchild prépare pour l'Army Electronics Command, une rétine contenant 1 000 éléments photosensibles, destinée à un lecteur automatique de pages.

Le SCT de la RCA diffère des CCD des Bell Laboratories par la présence d'une couche isolante, entre les électrodes et le semiconducteur. Son développement paraît compromis après l'invention des CCD à canal enterré.

REGISTRES A DÉCALAGE

Dispositifs	Puissance dissipée dans une cellule de registre	Cadence de fonctionnement	Surface occupée par une cellule de registre
Chaines à seaux	5 μ W	10 à 20 MHz	1 200 à 2 500 μ m ²
C.C.D.	5 μ W	10 à 1 000 MHz	1 200 à 2 500 μ m ²
Registres à transistors MOS classiques	100 μ W	5 MHz	3 750 à 7 500 μ m ²

LA COMPRESSION DES DONNÉES ET LEUR RECONNAISSANCE

La plupart des machines à lire dissectent chaque caractère en segments élémentaires qui sont ensuite corrélés dans l'unité logique. Cette segmentation est immédiate avec les matrices monolithiques ou discrètes de cellules photosensibles ; elle est nécessaire dans la lecture par flying-spot pour quantifier l'information.

Les techniques de segmentation sont nombreuses. Dans toutes les techniques, le dispositif de balayage doit initialement rechercher la première ligne d'une page à lire, par un balayage vertical de la page jusqu'à obtention de zones noires. Dès qu'une ligne est détectée, le lecteur balaie chaque caractère, découpé en petites cellules : dans certaines méthodes, jusqu'à 200 cellules peuvent être analysées par caractères : selon que la couleur prédominante d'une cellule est le blanc ou le noir, la machine lui affecte un 0 ou un 1 logique.

Chaque caractère se trouve ainsi représenté par une suite de 200 bits noirs ou blancs. Pour décrire le caractère par le code ASCII, il est nécessaire d'effectuer une « compression de données » en trouvant les 8 bits les plus caractéristiques.

Une variante de cette technique consiste à représenter un caractère par un code figurant le nombre et la position des lignes horizontales ou verticales qui le composent.

Une seconde variante consiste à représenter les caractères par ses courbures, par la direction des lignes qui le composent, par le nombre de lignes brisées, etc. Cette technique s'adapte à la reconnaissance de l'écriture manuscrite. Le balayage d'un caractère s'effectue suivant un mouvement circulaire ou suivant des lignes brisées ; ce balayage se fait selon les ordres donnés par une unité logique de reconnaissance. Cette méthode de lecture, étudiée pour l'IBM 1287, permet d'être utilisée pour l'écriture manuscrite.

Enfin, certains procédés utilisent directement des masques ou silhouettes : tous les caractères alphanumériques sont enregistrés sous forme codée dans la mémoire de l'unité de reconnaissance ; le caractère lu est comparé aux caractères enregistrés, afin de déterminer celui qui lui correspond. Les procédés purement optiques de corrélation par hologrammes se classent dans cette classe d'unités de reconnaissance : ces procédés sont très simples et puissants, mais manquent malheureusement encore de souplesse. Néanmoins, comparés aux procédés logiques, les procédés holographiques sont infiniment moins coûteux.

OU INTERVIENT L'HOLOGRAPHIE...

L'holographie est une technique basée sur les propriétés des interférences entre ondes, qu'elles soient lumineuses, radio-électriques ou acoustiques. Si les expériences de base de l'holographie remontent à 1948, la mise en œuvre pratique résulte de l'apparition du laser : c'est dans les années 1961 à 1964 que plusieurs chercheurs (Leith, Upatnieks, Van Der Lugt, Stroke, Denysik, etc.) révèlent toute la richesse des possibilités de l'holographie, en particulier de l'holographie optique.

Un hologramme est un enregistrement de phénomènes d'interférences produits par la superposition à une onde bien définie, la référence, des ondes cohérentes diffusées par l'objet à « holographier », lui-même éclairé par un faisceau d'ondes cohérentes.

Un objet quelconque, recevant de la lumière issue d'un laser, c'est-à-dire parfaitement ordonnée, de longueur d'onde bien définie, diffuse et diffracte dans toutes les directions. A un instant donné, le front d'onde venant de l'objet en constitue une véritable représentation : un observateur qui pourrait percevoir cette onde constaterait que son amplitude varie d'un point à un autre, mais

aussi que sa forme évolue, épousant les contours de l'objet.

L'holographie permet d'enregistrer, puis de restituer, la configuration et l'amplitude de ce signal caractéristique de la forme de l'objet. L'étape des restitutions consiste à éclairer, par le faisceau de référence avant servi lors de l'enregistrement, la plaque holographique. L'une des ondes diffractées possède exactement le profil de l'onde émise par l'objet lors de son enregistrement. L'image de l'objet apparaît en relief au travers de la fenêtre formée par le contour de l'hologramme.

...ET LE FILTRAGE DES FREQUENCES SPATIALES

Un signal sonore ou électrique est décomposable en série ou en intégrale de Fourier, c'est-à-dire en somme de termes, sinusoïdaux dont les amplitudes, les phases et les fréquences temporelles sont caractéristiques du signal. L'ensemble de ces vibrations élémentaires rangées par ordre de fréquences croissantes, constitue le spectre de Fourier du signal.

De manière analogue, une image optique bidimensionnelle peut être considérée comme la superposition de motifs sinusoïdaux, mires élémentaires d'amplitudes, de pas (ou de fréquences spatiales) et d'orientations différentes. C'est le spectre optique, également bidimensionnel.

En agissant sur le spectre optique, par filtrage, il est possible d'agir sur l'image et de la transformer profondément. Le spectre et l'objet « holographié » (c'est le signal) sont liés par une transformation de Fourier, le spectre et l'image également. De ce fait, en agissant convenablement sur le spectre, on peut effectuer un certain nombre d'opérations mathématiques sur le signal. L'image ne sera plus une image au sens courant du terme, mais une réponse.

La possibilité d'identifier un signal par filtrage des fréquences spatiales a été démontrée par E. O'Neill, aux U.S.A., en 1956. Le filtre qu'il faut utiliser possède une loi de transmission complexe, pouvant être pratiquement réalisée par holographie.

L'application du filtrage des fréquences spatiales est développée, en France en particulier à l'Institut d'Optique de Paris, et au Laboratoire d'Optique de Besançon. Le principe de la reconnaissance est le suivant : le dispositif optique de reconnaissance, d'un texte par exemple, possède en mémoire, sous forme de filtre « adapté », le signal recherché. L'analyse s'effectue en produisant d'abord

le spectre du texte à reconnaître, en le plaçant devant une lentille : le spectre apparaît dans le plan focal ; en l'absence de toute opération, on obtiendrait, dans la plan image, une image renversée du texte. Le filtre adapté, placé dans le plan focal, fait disparaître cette image mais laisse passer les composantes spectrales de la lettre recherchée. La réponse apparaît, dans le plan image, sous la forme d'un point, dont la présence indique que la lettre a été identifiée, et dont la position révèle la position de la lettre. Le schéma d'une machine à lire, basée sur ce principe, paraît alors immédiat : en faisant défiler 26 filtres dans le plan focal, chaque filtre correspondant à une lettre de l'alphabet, on pourrait lire le texte entier. Les informations digitales obtenues (les points) peuvent ensuite être introduites, par un tube de caméra de télévision, dans une machine arithmétique relativement simple.

RECONNAITRE DES IMAGES

Si la machine holographique à lire n'existe pas encore, d'autres dispositifs, basés sur le filtrage des fréquences spatiales sont déjà réalisées. C'est ainsi que l'on traite des signaux géophysiques.

J. Duvernoy, de l'Université de Besançon a démontré, au Symposium international sur les applications de l'holographie (1970) qu'il est possible de traiter des coupes sismiques au moyen de filtres-hologrammes, afin d'en extraire l'information utile concernant la structure du sous-sol. La firme américaine Bendix a conçu, suivant ce principe, un appareillage capable de tracer les lignes isométriques sur une photographie aérienne, et, donc d'en extraire des informations sur le relief de la région photographiée.

On imagine aisément la diversité des autres applications offertes à la reconnaissance optique des formes, qu'il s'agisse de la détermination d'une trajectoire précise dans une chambre à bulles ou des empreintes digitales d'un individu dans un fichier qui en contient plusieurs milliers ; ce dernier domaine pourrait prendre de l'ampleur : KMS Industries Inc. a livré des appareillages de reconnaissance des empreintes digitales à l'U.S. Force, et également à une banque de New York ; chaque unité coûte 7 500 dollars. McDonnell-Douglas Electronics Co. prépare un modèle beaucoup plus élaboré, au prix de 250 000 dollars, dont le marché potentiel a été évalué à 350 unités.

Marc FERRETTI