(19) RÉPUBLIQUE FRANÇAISE

INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE

PARIS

(11) N° de publication : **2 661 999**
(à n'utiliser que pour les commandes de reproduction)

(21) N° d'enregistrement national : **90 06081**

(51) Int Cl[5] : G 06 K 9/68

(12) DEMANDE DE BREVET D'INVENTION A1

(22) **Date de dépôt** : 11.05.90.

(30) **Priorité** :

(43) **Date de la mise à disposition du public de la demande :** 15.11.91 Bulletin 91/46.

(56) **Liste des documents cités dans le rapport de recherche :** *Se reporter à la fin du présent fascicule.*

(60) **Références à d'autres documents nationaux apparentés :**

(71) **Demandeur(s) :** *UNIVERSITE DE HAUTE ALSACE (Etablissement public à caractère scientifique et culturel)* — FR.

(72) **Inventeur(s) :** Metzger Gérard.

(73) **Titulaire(s) :**

(74) **Mandataire :** Cabinet Nithardt & Burkard S.A.

(54) Procédé et dispositif d'identification de formes ou de profils.

(57) La présente invention concerne un procédé et un dispositif pour l'identification de formes et de profils. Selon ce procédé on crée par une analyse d'images, un ensemble de données informatiques correspondant à l'image d'une forme ou d'un profil à identifier ou à authentifier.

Ce procédé est caractérisé en ce que l'on crée ledit fichier et ledit ensemble de données informatiques en un langage descriptif de l'image utilisant des courbes de Bézier et/ou des B et/ou C Spline.

Le procédé s'applique notamment à la reconnaissance de caractères typographiques.

FR 2 661 999 - A1

PROCEDE ET DISPOSITIF D'IDENTIFICATION DE FORMES OU DE PROFILS

La présente invention concerne un procédé d'identification de formes ou de profils notamment pour la reconnaisance de caractères typographiques, dans lequel on crée par analyse d'images un ensemble de données informatiques correspondant à l'image d'une forme ou d'un profil à identifier.

L'invention concerne également un dispositif d'identification de formes ou de profils, notamment pour la reconnaissance de caractères typographiques, comportant des moyens pour créer, par analyse d'images, un ensemble de données informatiques correspondant à l'image d'une forme ou d'un profil à identifier pour la mise en oeuvre du procédé ci-dessus.

La présente invention s'applique plus particulièrement à l'identification de formes telles que par exemple des caractères de machine à écrire, en vue d'une authentification ou de l'identification d'une machine à écrire.

On connaît un certain nombre de systèmes permettant d'atteindre les buts recherchés mentionnés précédemment. Dans la plupart des procédés connus, un ordinateur contient en mémoire la représentation sous forme d'une matrice de points des formes à reconnaître ou un certain nombre de formes types qui constituent un certain nombre de données informatiques obtenues par analyse d'images, par exemple au moyen d'un système de balayage. La matrice de points forme une image en noir et blanc où les points noirs représentent par exemple un profil à reconnaître ou à identifier. La reconnaissance est faite par une corrélation mathématique entre la forme type, c'est-à-dire celle qui est enregistrée dans l'ordinateur, et une forme lue, c'est-à-dire celle qui est transmise par un système d'analyse d'images et qui correspond à une forme à identifier.

La reconnaissance est d'autant plus précise qu'il y a plus de points dans la représentation. Il en résulte que la qualité de la reconnaissance dépend particulièrement de la résolution. De ce fait, pour obtenir un travail précis et fiable, il est nécessaire de travailler avec un matériel à

grande résolution et par conséquent avec des ordinateurs de grande capacité.

Ce type d'ordinateurs est relativement coûteux. D'autre part, le temps d'analyse et d'authentification est d'autant plus long que le nombre de points à vérifier est important, c'est-à-dire que la résolution du système est grande, ou que le nombre de caractères à analyser est grand.

En conclusion, avec les systèmes connus, si l'on souhaite pouvoir disposer de systèmes fiables, il est nécessaire de disposer d'un équipement coûteux. En outre, dans ce cas, l'analyse des informations est longue ce qui constitue un handicap majeur étant donné qu'un système efficace va pouvoir effectuer une identification dans un laps de temps très court.

La présente invention se propose de pallier les inconvénients des systèmes connus et d'offrir à la fois un procédé et un dispositif d'identification très efficace, très rapide, et travaillant avec un matériel informatique standard.

Dans ce but, le procédé selon l'invention est caractérisé en ce que l'on crée ledit ensemble de données informatiques en un langage descriptif de l'image, utilisant des courbes de Bézier et/ou des B et/ou des C Spline.

Selon une forme de réalisation préférée, ledit langage descriptif est du type Postscript.

Selon un mode de réalisation particulièrement avantageux, pour créer lesdites données informatiques dans ledit langage descriptif de l'image, on définit le contour de la forme à identifier au moyen de courbes de Bézier, et on détermine les coefficients des polynômes de Bézier correspondant à ces courbes.

Pour identifier une forme ou un profil, on compare avantageusement les coefficients de Bézier des polynômes de Bézier correspondant aux coefficients de Bézier des polynômes de Bézier mémorisés pour

constituer un fichier.

Selon une forme de réalisation préférée, on effectue cette comparaison par corrélation numérique pondérée pour tenir compte de transformations géométriques simples et des distorsions de la lecture numérique.

Le dispositif pour la mise en oeuvre de ce procédé est caractérisé en ce qu'il comporte des moyens pour créer ledit ensemble de données informatiques dans ledit langage descriptif de l'image.

Selon un mode de réalisation préféré, lesdits moyens sont agencés pour travailler dans un langage du type Postscript.

D'une manière avantageuse, le dispositif comporte lesdits moyens pour créer lesdites données informatiques dans ledit langage descriptif de l'image et sont agencées pour définir le contour de la forme à identifier au moyen de courbes de Bézier et pour déterminer les coefficients des polynômes de Bézier correspondant à ces courbes.

Selon un mode de réalisation préféré, il comporte des moyens pour comparer les coefficients de Bézier des polynômes de Bézier correspondant à la forme ou au profil à identifier aux coefficients de Bézier des polynômes de Bézier mémorisés pour constituer un fichier.

Selon une forme de réalisation particulièrement avantageuse, le dispositif de corrélation numérique pondérée est agencé pour tenir compte de transformations géométriques simples et de distorsion de lecture numérique.

La présente invention sera mieux comprise en référence à la description d'une forme de réalisation préférée et de variantes avantageuses permettant d'atteindre les buts précisés ci-dessus.

Dans le cas du procédé, une première phase de l'identification consiste à créer un fichier de données informatiques correspondant à un ensemble de formes ou de profils standards connus auxquels on envisage de

comparer une forme ou un profil à identifier. La démarche pour créer le fichier et celle consistant à créer un ensemble de données informatiques correspondant à l'image d'une forme ou d'un profil à identifier sont strictement identiques. Cette phase appelée définition ou création de la forme est de préférence effectuée au moyen d'un logiciel de décalque Postscript fonctionnant sur un ordinateur standard du type mini-ordinateur tel que les stations commercialisées sous la dénomination "PC" de la Société IBM ou ordinateur compatible ou sur une station du type commercialisée par la Société MacIntosh. On obtient un fichier descriptif des formes, par exemple en format normalisé EPSF ou I 118.

Le principe consiste à obtenir les coefficients des polynômes de Bézier des courbes de Bézier du contour de chaque forme. Pour cela, les courbe de Bézier sont ajustées au contour par un suivi du contour. Chaque image est convertie en un ensemble de formes, dont chacune est définie par un chemin de Bézier lui-même déterminé par un ensemble de coefficients. La condensation de l'image est maximale, étant donné qu'une liste de coefficients définit tout le contenu de cette image.

La deuxième phase, qui correspond à la reconnaissance proprement dite, consiste en une comparaison des coefficients de Bézier d'une forme type prise dans le fichier mémorisé avec les coefficients de Bézier d'une forme à identifier. Cette comparaison s'effectue de préférence par corrélation numérique pondérée pour tenir compte de transformations géométriques élémentaires telles que des rotations, des modifications de dimensions, des translations (par exemple par une transformation de Hough).

Comme mentionné précédemment, le langage descriptif présente l'avantage d'être extrêmement concis tout en fournissant les données informatiques correspondant à une définition très précise d'une forme.

La transmission de fichiers PostScript EPSF ou I 118, ou uniquement les coefficients numériques peuvent être transmis par les lignes téléphoniques, ce qui permet de réaliser à distance la description d'une image. Etant donné qu'en langage descriptif de l'image, la masse des données

nécessaires pour décrire une image est relativement réduite, cette transcription peut s'effectuer très rapidement.

Pour illustre un mode d'utilisation du dispositif ci-dessus, on se propose de décrire un dispositif d'identification-authentification d'une machine à écrire. Dans cet exemple, la police dispose par exemple d'une bibliothèque de caractères types, frappés par des machines diverses dont certaines sont encore disponibles et d'autres non. Lorsqu'elle est en présence d'un courrier écrit par une machine à déterminer, on peut utiliser le dispositif ci-dessus. Ce dispositif permet d'effectuer l'analyse successive de tous les caractères du courrier et d'en faire une description dans ledit langage informatique. Il en résulte une liste de coefficients pour chaque caractère à authentifier. Une corrélation numérique pondérée est alors effectuée avec les descriptions des mêmes caractères extraits de la bibliothèque. On obtient une probabilité de reconnaissance qui est maximale pour la machine authentique qui a frappé le courrier ou, si elle n'est pas disponible, pour la machine de même type et série de construction.

Les avantages de ce système sont essentiellement dus à la compacité des descriptions PostScript, c'est-à-dire au nombre réduit de coefficients de Bézier nécessaires pour définir une forme ou un profil. De ce fait, la reconnaissance est rapide et ne nécessite qu'un ordinateur courant et utilisant des logiciels standards. La qualité de la détection est excellente de par l'utilisation d'un langage descriptif de l'image qui permet la description de toutes les formes et aussi d'effectuer toute transformation géométrique sur cette image d'une manière extrêmement simple.

La présente invention n'est pas limitée aux formes de réalisations décrites mais peut s'étendre à différentes variantes évidentes pour l'homme de l'art et peut subir diverses modifications, sans pour autant dépasser le cadre de cette invention défini par les revendications.

Revendications

1. Procédé d'identification de formes ou profils, notamment pour la reconnaissance de caractères typographiques, dans lequel on crée par une analyse d'images un ensemble de données informatiques correspondant à l'image d'une forme ou d'un profil à identifier, caractérisé en ce que l'on crée ledit ensemble de données informatiques en un langage descriptif de l'image, utilisant des courbes de Bézier et/ou des B et/ou des C Spline.

2. Procédé selon la revendication 1, caractérisé en ce que ledit langage descriptif est du type PostScript.

3. Procédé selon la revendication 2, caractérisé en ce que pour créer lesdites données informatiques dans ledit langage descriptif de l'image, on définit le contour de la forme à identifier au moyen de courbes de Bézier et on détermine les coefficients des polynômes de Bézier correspondant à ces courbes.

4. Procédé selon la revendication 3, caractérisé en ce que pour identifier une forme et un profil, on compare les coefficients de Bézier des polynômes de Bézier correspondant aux coefficients de Bézier des polynômes de Bézier mémorisés pour constituer un fichier.

5. Procédé selon la revendication 4, caractérisé en ce que l'on effectue cette comparaison par corrélation numérique pondérée pour tenir compte de transformations géométriques simples, et des distorsions de la lecture numérique.

6. Dispositif d'identification de formes ou de profils, notamment pour la reconnaissance de caractères typographiques, comportant des moyens pour créer, par analyse d'images, un ensemble de données informatiques correspondant à l'image d'une forme ou d'un profil à identifier, caractérisé en ce qu'il comporte des moyens pour créer ledit ensemble de données informatiques dans ledit langage descriptif de l'image.

7. Dispositif selon la revendication 6, caractérisé en ce que lesdits moyens sont agencés pour travailler dans un langage du type Postscript.

8. Dispositif selon la revendication 6, caractérisé en ce que lesdits moyens pour créer lesdites données informatiques dans ledit langage descriptif de l'image, sont agencés pour définir le contour de la forme à identifier au moyen de courbes de Bézier et pour déterminer les coefficients des polynômes de Bézier correspondant à ces courbes.

9. Dispositif selon la revendication 8, caractérisé en ce qu'il comporte des moyens pour comparer les coefficients de Bézier des polynômes de Bézier correspondant à la forme ou au profil à identifier aux coefficients de Bézier des polynômes de Bézier mémorisés pour constituer un fichier.

10. Dispositif selon la revendication 9, caractérisé en ce que les moyens par comparaison comportent un dispositif de corrélation numérique pondérée agencé pour tenir compte de transformations géométriques simples, et de distorsions de lecture numérique.

REPUBLIQUE FRANÇAISE

INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLE

2661999

N° d'enregistrement national

RAPPORT DE RECHERCHE

établi sur la base des dernières revendications déposées avant le commencement de la recherche

FR 9006081
FA 442168

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendications concernées de la demande examinée
A	PATTERN RECOGNITION. vol. 21, no. 1, 1988, US pages 9 - 17; IWAO SEKITA ET AL.: "FEATURE EXTRACTION OF HANDWRITTEN JAPANESE CHARACTERS BY SPLINE FUNCTIONS FOR RELAXATION MATCHING" * page 10, colonne de gauche, alinéa 2 - colonne de droite, alinéa 1 *	1
A	COMPUTER VISION GRAPHICS AND IMAGE PROCESSING. vol. 39, no. 3, septembre 1987, NEW YORK US pages 267 - 278; G. MEDIONI ET AL.: "CORNER DETECTION AND CURVE REPRESENTATION USING B-SPLINES" * abrégé *	1
A	EP-A-163885 (SIEMENS AG) * page 2, lignes 5 - 25 *	1
A	17TH ANNUAL ACM COMPUTER SCIENCE CONFERENCE, NEW YORK 1989 pages 162 - 165; Wm. E. Ellis: "POSTSCRIPT, BEZIER CURVES AND CHINESE CHARACTERS"	

DOMAINES TECHNIQUES RECHERCHES (Int. Cl.5)

G06K9/00

Date d'achèvement de la recherche	Examinateur
22 JANVIER 1991	SONIUS M.E.

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général
O : divulgation non-écrite
P : document intercalaire

T : théorie ou principe à la base de l'invention
E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure.
D : cité dans la demande
L : cité pour d'autres raisons

& : membre de la même famille, document correspondant

EPO FORM 1503 03.82 (P0413)

1