

Novática, revista fundada en 1975 y decana de la prensa informática española, es el órgano oficial de expresión y formación continua de ATI (Asociación de Técnicos de Informática), organización que edita también la revista REICIS (Revista Española de Innovación, Calidad e Ingeniería del Software). **Novática** edita asimismo UPGRADE, revista digital de CEPIS (Council of European Professional Informatics Societies), en lengua inglesa, y es miembro fundador de UPENET (UPGRADE European NETWORK).

<<http://www.ati.es/novatica/>>
<<http://www.ati.es/reicis/>>
<<http://www.upgrade-cepis.org/>>

ATI es miembro fundador de CEPIS (Council of European Professional Informatics Societies) y es representante de España en IFIP (International Federation for Information Processing); tiene un acuerdo de colaboración con ACM (Association for Computing Machinery), así como acuerdos de vinculación o colaboración con AdaSpain, AIZ, ASTIC, RITSI e Hispalinux, junto a la que participa en Prolnova.

Consejo Editorial

Antoni Carbonell Nogueras, Juan Manuel Cueva Lovelle, Juan Antonio Esteban Iriarte, Francisco López Crespo, Julián Marcelo Cocho, Celestino Martín Alonso, Josep Molas i Bertrán, Olga Pallas Cerdina, Fernando Píera Gomez (Presidente del Consejo), Ramón Puigjaner Trepal, Miquel Sàrries Grifó, Asunción Yturbe Herranz

Coordinación Editorial

Llorenç Pagés Casas <lpages@ati.es>

Composición y autoedición

Jorge Llácer Gil de Rameles

Traducciones

Grupo de Lengua e Informática de ATI <<http://www.ati.es/gt/lengua-informatica/>>, Dpto. de Sistemas Informáticos - Escuela Superior Politécnica - Universidad Europea de Madrid

Administración

Tomás Brunete, María José Fernández, Enric Camarero, Felicidad López

Secciones Técnicas - Coordinadores

Acceso y recuperación de la información

José María Gómez Hidalgo (Universidad Europea de Madrid), <jmgomez@uem.es>

Manuel J. Mañá López (Universidad de Huelva), <manuel.mana@dieia.uhu.es>

Administración Pública electrónica

Francisco López Crespo (MAE), <flc@ati.es>

Gumersindo García Arribas (MAP), <gumersindo.garcia@map.es>

Arquitecturas

Enrique F. Torres Moreno (Universidad de Zaragoza), <enrique.torres@unizar.es>

Jordi Tubellà Morgadas (DAC-UPC), <jordit@ac.upc.es>

Auditoría OTIC

Marina Tourinho Troitiño, <marinatourino@marinatourino.com>

Manuel Palao García-Suelto (ASIA), <manuel@palao.com>

Derecho y tecnologías

Isabel Hernando Colino (Fac. Derecho de Donostia, UPV), <ihernando@legalek.net>

Elena Davara Fernández de Marcos (Davara & Davara), <edavara@davara.com>

Educación Universitaria de la Informática

Joaquín Ezpeleta Mateo (OPS-UZAR), <ezpeleta@posta.unizar.es>

Cristóbal Pavia Flores (DSIP-UCM), <cpavia@sis.ucm.es>

Entorno digital personal

Alonso Alvarez García (TID), <aag@tid.es>

Diego Gachet Páez (Universidad Europea de Madrid), <gachet@uem.es>

Estándares Web

Encarnación Duesada Ruiz (Oficina Española del W3C), <eduesada@w3.org>

José Carlos del Arco Prieto (TCP Sistemas e Ingeniería), <jarco@gmail.com>

Gestión del Conocimiento

Juan Baiget Solé (Cap Gemini Ernst & Young), <joan.baiget@ati.es>

Informática y Filosofía

José Ángel Olivas Varela (Escuela Superior de Informática, UCLM)

Karim Gherab Martín (Indra Sistemas)

Informáticos Gráficos

Miguel Chover Saltes (Universitat Jaume I de Castellón), <chover@lsi.uji.es>

Roberto Vivó Hernando (Eurographics, sección española), <rvivo@dsic.upv.es>

Ingeniería del Software

Javier Dolado Costin (D.S.I-UPV), <dolado@si.uhu.es>

Luis Fernández Sanz (PRIS-El-UEM), <lufern@dpriis.esi.uem.es>

Inteligencia Artificial

Vicente Botti Navarro, Vicente Julián Inglada (DSIC-UPV)

<vbotti_vinjadela@dsic.upv.es>

Interacción Persona-Computador

Julio Abascal González (FI-UPV), <julio@si.uhu.es>

Lengua e Informática

M. del Carmen Ugarte García (IBM), <cugarte@ati.es>

Lenguajes Informáticos

Andrés Marín López (Univ. Carlos III), <amarin@it.uc3m.es>

J. Angel Velázquez Turbide (ESDET-URJC), <a.velazquez@esdet.urjc.es>

Limpieza computacional

Xavier Gómez Guinovart (Univ. de Vigo), <xgg@uvigo.es>

Manuel Palomar (Univ. de Alicante), <mpalomar@disi.ua.es>

Mundo estudiantil y jóvenes profesionales

Federico G. Mon Troiti (RITSI) <gmon@ati.es>

Mikel Salazar Páez (Araiz), Jóvenes Profesionales, Junta de ATI Madrid), <mikeltxo_uni@yahoo.es>

Profesión Informática

Rafael Fernández Calvo (ATI), <rflcalvo@ati.es>

Miquel Sàrries Grifó (Ayto. de Barcelona), <misarries@ati.es>

Redes y servicios telemáticos

José Luis Marzo Lázaro (Univ. de Girona), <jose.luis.marzo@udg.es>

Josép Solé Pareta (DAC-UPC), <pareta@ac.upc.es>

Seguridad

Javier Arellito Bertolin (Univ. de Deusto), <jarellito@eside.deusto.es>

Javier López Muñoz (ETSI Informática-UMA), <jlm@cc.uma.es>

Sistemas de Tiempo Real

Alejandro Alonso Muñoz, Juan Antonio de la Puente Alfaro (DIT-UPM),

<[@dit.upm.es">alalons@lpuente@dit.upm.es](mailto:alalons@lpuente)>

Software Libre

Jesús M. González Barahona, Pedro de las Heras Quirós (GSYC-URJC),

<[@gsyc.urjc.es">jgib@heras@gsyc.urjc.es](mailto:jgib@heras)>

Tecnología de Objetos

Jesús García Molina (DIS-UM), <jgmolina@um.es>

Gustavo Rossi (LIFA-UNLP, Argentina), <gustavo@sol.inf.unlp.edu.ar>

Tecnologías para la Educación

Juan Manuel Dodero Beardo (UC3M), <ddodero@inf.uc3m.es>

Juliá Minguillón i Alfonso (UOC), <jminguillona@uoc.edu>

Tecnologías y Empresa

Didac López Dufuill (Universitat de Girona), <didac.lopez@ati.es>

Francisco Javier Cantais Sánchez (Indra Sistemas), <fcantais@gmail.com>

TIC y Turismo

Andrés Aguayo Maldonado, Antonio Guevara Plaza (Univ. de Málaga)

<[@lcc.uma.es">aguayo.guevara@lcc.uma.es](mailto:aguayo.guevara)>

Las opiniones expresadas por los autores son responsabilidad exclusiva de los mismos.

Novática permite la reproducción, sin ánimo de lucro, de todos los artículos, a menos que lo impida la modalidad de © o copyright elegida por el autor, debiéndose en todo caso citar su procedencia y enviar a **Novática** un ejemplar de la publicación.

Coordinación Editorial, Redacción Central y Redacción ATI Madrid

Padilla 66, 3º dcha., 28006 Madrid

Tel. 91 4029391 - fax 91 3093685 <novatica@ati.es>

Composición, Edición y Redacción ATI Valencia

Av. del Reino de Valencia 23, 46005 Valencia

Tel./fax 963309392 <secretval@ati.es>

Administración y Redacción ATI Cataluña

Via Llorens 46, local 1º, 08003 Barcelona

Tel. 934125235 - fax 934127713 <secretgen@ati.es>

Redacción ATI Andalucía

Isaac Newton, s/n, Ed. Sadies

Islas Cartuja 41092 Sevilla, Tel./fax 954460779 <secretand@ati.es>

Redacción ATI Aragón

Lagasca 9, 3-B, 50006 Zaragoza,

Tel./fax 978235181 <secretara@ati.es>

Redacción ATI Asturias-Cantabria

Redacción ATI Castilla-La Mancha <gp-clmancha@ati.es>

Sección de Ventas

<<http://www.ati.es/novatica/interes.html>>, ATI Cataluña, ATI Madrid

Publicidad

Padilla 66, 3º dcha., 28006 Madrid

Tel. 91 4029391 - fax 91 3093685 <novatica@ati.es>

Impreso: Heras S.A., Juan de Austria 66, 08005 Barcelona

Depósito legal: B 15.154-1975 -- ISSN: 0211-2124; CODEN NOVACB

Partada: The white cybernautical cane, Concha Arias Pérez / © ATI 2007

Diseño: Fernando Agresta / © ATI 2007

Nº 186, marzo-abril 2007, año XXXIII

sumario

editorial

Nuestra formación como profesionales europeos

> 02

en resumen

Hacia una igualdad de oportunidades

> 02

Llorenç Pagés Casas

noticias IFIP

Aires de renovación en IFIP

> 03

Ramón Puigjaner Trepal

monografía

Informática para discapacitados visuales

(En colaboración con UPGRADE)

Editores invitados: *Josep Lladós Canet, Jaime López Krahe y Dominique Archambault*

Presentación: introducción a las ayudas técnicas para las personas ciegas

> 04

Jaime López Krahe

Informática a ciegas

> 08

Carmen Bonet Borrás

Tecnología y Discapacidad Visual

> 12

José Antonio Muñoz Sevilla

Visión general de las tecnologías de ordenadores de mano para

personas con disfunciones visuales

> 16

Philippe Foucher

El acceso a contenido científico de personas con discapacidad visual

> 23

Dominique Archambault, Bernhard Stöger, Donal Fitzpatrick, Klaus Miesenberger

Juegos de ordenador y discapacitados visuales

> 28

Dominique Archambault, Roland Ossmann, Thomas Gaudy, Klaus Miesenberger

Herramientas de Visión por Computador para el aprendizaje de niños invidentes

> 33

Gemma Sánchez Albaladejo, Alicia Fornés Bisquerra, Joan Mas Romeu, Josep Lladós Canet

Tecnología y educación en el campo de la discapacidad visual

> 39

Silvia Boix Hernández, M^a Teresa Corbella Roqueta, Lucía Melchor Sánchez

SAW: un conjunto de herramientas integradas para hacer la web accesible

> 43

a usuarios invidentes

Fernando Sánchez Figuerola, Adolfo Lozano-Tello, Julia González Rodríguez, Mercedes

Macías García

Adaptación a la normativa WAI de manera automática

> 47

Juan Manuel Fernández Ramirez, Vicenç Soler Ruiz, Jordi Roig de Zárata

Ayudando a los autores a la generación de contenido accesible:

dos experiencias europeas

> 52

Carlos Rebate Sánchez, Alicia Fernández del Viso Torre

secciones técnicas

Enseñanza Universitaria de la Informática

Perfiles de egresado en las titulaciones de Informática y medios

> 56

para el empleo de las universidades españolas

Luis Fernández Sanz, María José García García, Verónica Egido García, Mario Mata Ortega,

Nouridine Aliane Saadi

Mundo estudiantil y jóvenes profesionales

Vaelsys: visión artificial a través de un proyecto emprendedor

> 61

Eduardo Cermeño Mediavilla, Carlos Jesús Venegas Arrabé

Referencias autorizadas

> 65

sociedad de la información

Informática e integración social

La evolución hacia una nueva brecha digital

> 71

Evelio Martínez Martínez, Arturo Serrano Santoyo

Programar es crear

Saldando cuentas (CUPCAM 2006, problema C, enunciado)

> 75

Manuel Carro Linares, Manuel Freire Morán

El juego de los engranajes (CUPCAM 2006, problema B, solución)

> 76

Manuel Abellanas Oar, Manuel Freire Morán

asuntos interiores

Coordinación editorial / Socios Institucionales

> 77

Monografía del próximo número: "Certificaciones profesionales en Informática"

Gemma Sánchez Albaladejo,
Alicia Fornés Bisquerra, Joan
Mas Romeu, Josep Lladós
Canet

Centre de Visió per Computador, Departament
de Ciències de la Computació, Universitat
Autònoma de Barcelona

{gemma, afornes, jmas, josep}@cvc.uab.cat

Herramientas de Visión por Computador para el aprendizaje de niños invidentes

1. Introducción

La visión por computador es una tecnología que estudia la manera de extraer información de las imágenes, mediante un ordenador, para poder comprenderlas. Uno de los campos de la visión por computador es el análisis de documentos, que combina técnicas de análisis de imágenes y de reconocimiento de formas para procesar e interpretar documentos electrónicos [1]. Estos documentos pueden estar en papel y ser digitalizados mediante un escáner o estar ya en un formato electrónico como el pdf. Esta línea de investigación se divide en dos grandes subáreas de interés: el reconocimiento de información textual (en inglés OCR *Optical Character Recognition*) y el reconocimiento de información gráfica como la que aparece en planos, dibujos, mapas, etc. El OCR es una de las actividades pioneras del reconocimiento de formas y se pueden encontrar trabajos en esta línea desde los años 50 [2]. El reconocimiento gráfico, en cambio, puede considerarse una actividad más reciente que ha tenido un crecimiento importante en los últimos tiempos debido inicialmente a la necesidad de digitalizar planos y mapas para introducirlos de forma automática en entornos CAD y GIS, y últimamente a la aparición de dispositivos basados en la interacción manuscrita como los ordenadores de bolsillo (PDA's) [3] [4].

La visión por computador tiene un creciente interés como tecnología de soporte a las personas invidentes. Ejemplos como lectores de códigos de barras para identificar productos en supermercados, identificadores de colores para elegir prendas de ropa, o lectores de documentos, son aplicaciones que, combinadas a mecanismos de síntesis de voz, son herramientas cotidianas. Este artículo describe dos aplicaciones de visión por computador destinadas a niños invidentes y a sus profesores¹. Estas aplicaciones están en fase de desarrollo, y se están realizando en el Centro de Visión por Computador (CVC) de la Universitat Autònoma de Barcelona (UAB) con la colaboración del Centro de Recursos Educativos Joan Amades de la ONCE. La primera consiste en un traductor de Braille a *vista*. *Vista* es la forma en que se denomina a la escritura sobre papel impreso en tinta a la que estamos acostum-

Resumen: *la Visión por Computador es un área de la Inteligencia Artificial que se centra en la definición de modelos cognitivos artificiales para el procesamiento visual de la información. Puede utilizarse para desarrollar herramientas útiles para las personas invidentes ya que, con ayuda de un sensor óptico como una cámara digital, permite extraer la información asociada a estímulos visuales. En este artículo presentamos dos herramientas software desarrolladas para niños invidentes o con baja visión. La primera es un sistema de traducción de textos Braille escaneados. La segunda aplica técnicas de reconocimiento dinámico de formas para interpretar figuras geométricas sencillas dibujadas sobre una tableta digitalizadora. Esta herramienta permite a los niños aprender, de forma autónoma, conceptos básicos de forma o relaciones espaciales.*

Palabras clave: *análisis de documentos, interfaces caligráficas, reconocimiento de gráficos.*

Autores

Gemma Sánchez Albaladejo obtuvo el título de Ingeniería Superior en Informática en 1994 por la *Universitat Politècnica de Catalunya* y el título de Doctora en Informática en 2001 por la *Universitat Autònoma de Barcelona* y la *Université Henry Poincaré* (Francia). Actualmente es profesora en el departamento de Ciencias de la Computación de la *Universitat Autònoma de Barcelona* y miembro del Consejo de Dirección del Centro de Visión por Computador donde ha participado en diversos proyectos industriales y de investigación. Sus principales áreas de interés en investigación son el análisis de documentos, el reconocimiento de gráficos, y el reconocimiento de patrones sintácticos y estructurales. Actualmente trabaja en el dominio de los lenguajes visuales para la representación y reconocimiento de gráficos y *word spotting*. Ha publicado diversos artículos en conferencias nacionales e internacionales y en revistas. Es miembro del TC-10 IAPR (*Technical Committee on Graphics Recognition*) y editora adjunta de la revista electrónica internacional ELCVIA (*Electronic Letters on Computer Vision and Image Analysis*).

Alicia Fornés Bisquerra obtuvo el título de Ingeniera Informática en 2003 por la *Universitat de les Illes Balears* y en 2005 obtuvo el Diploma de Estudios Avanzados (DEA-tesina) por la *Universitat Autònoma de Barcelona*. Actualmente está realizando el doctorado en visión por computador y trabaja como profesora ayudante en la *Universitat Autònoma de Barcelona*. Sus principales áreas de interés son el análisis de documentos, el reconocimiento de gráficos y antiguos manuscritos y el reconocimiento óptico musical. Ha publicado varios artículos en conferencias nacionales e internacionales.

Joan Mas Romeu obtuvo el título de Ingeniería Informática en el año 2003 y el título de Master en Visión por Computador en el año 2005, ambos por la *Universitat Autònoma de Barcelona*. Actualmente trabaja como profesor ayudante y realiza su tesis doctoral en la *Universitat Autònoma de Barcelona*. Sus líneas de investigación son: el análisis de documentos, reconocimiento gráfico y reconocimiento estructural y sintáctico de patrones. Actualmente trabaja en el desarrollo de lenguajes visuales para la representación y reconocimiento de documentos gráficos. Ha publicado varios artículos en conferencias nacionales e internacionales.

José Lladós Canet obtuvo el título de Licenciado en Informática en 1991 por la *Universitat Politècnica de Catalunya* y el doctorado en Informática en 1997 por la *Universitat Autònoma de Barcelona* y la *Université Paris 8*. Actualmente es Catedrático de Escuela Universitaria del Depto. de Ciencias de la Computación de la UAB e investigador del *Centre de Visió per Computador* (CVC), del que es además el subdirector. Es el director del grupo de Reconocimiento de Formas y Análisis de Documentos (2005SGR-00472). Sus líneas de investigación son el análisis de documentos y el reconocimiento estructural y sintáctico de formas. Ha pilotado diversos proyectos de I+D en Visión por Computador y publicado más de 50 artículos en congresos y revistas especializados. Josep Lladós es miembro de la *Asociación Española de Reconocimiento de Formas y Análisis de Imágenes* (AERFAI), delegación española de la IAPR (*International Association for Pattern Recognition*). Actualmente es el responsable del IAPR-ILC (*Industrial Liaison Committee*). Anteriormente fue coordinador del IAPR TC-10 (*Technical Committee on Graphics Recognition*). Es miembro del comité editorial de las revistas ELCVIA (*Electronic Letters on Computer Vision and Image Analysis*) e IJDAR (*International Journal in Document Analysis and Recognition*), y miembro del comité de programa de diferentes congresos internacionales. Cuenta también con experiencia en transferencia de tecnología y en 2002 fundó la compañía ICAR Vision Systems, como spin-off del CVC.



Figura 1. a) Sistema de traducción Braille-Vista. b) Resultado de la traducción.

brados las personas no invidentes, y se utiliza en contraposición al término Braille para diferenciar ambas escrituras. Traducir textos Braille de forma automática permite a los niños compartir sus trabajos con los demás, independientemente de sus conocimientos de Braille, además de duplicar textos Braille sin necesidad de reescribirlos. La segunda aplicación permite a los niños invidentes y con resto visual aprender, de forma autónoma, a dibujar formas geométricas o letras mayúsculas y a reseguirlas utilizando una tableta digitalizadora. Este aprendizaje es útil en diferentes ámbitos. Por ejemplo, aprender a seguir líneas paralelas es un paso en el aprendizaje de la lectura Braille, ya que para realizarla el niño ha de seguir los caracteres Braille que aparecen dispuestos en líneas paralelas. Aprender a escribir pequeños textos también puede ayudarle a comunicarse con los demás o dejarles pequeñas notas. Por último, ser capaz de dibujar paisajes, el cuerpo humano o un esquema de cualquier aspecto de la realidad le permite poder concebir representaciones a escala de esta realidad y le ayuda a entender su disposición en el espacio. Además, dibujar a mano alzada formas predefinidas que el sistema identifique permite que éste las redibuje eliminando posibles imperfecciones y mejorando el resultado final de un documento.

Desde un punto de vista científico ambas actuaciones son aplicaciones de las dos grandes líneas de investigación del análisis de documentos nombradas anteriormente: análisis de texto y de gráficos. Así el objetivo de la traducción de textos Braille a vista es un caso particular de OCR llamado OBR

(*Optical Braille Recognition*) [5], y el reconocimiento interactivo de figuras gráficas diseñadas mediante una tableta digitalizadora es un caso de aplicación de técnicas de reconocimiento de símbolos gráficos [6].

A continuación detallamos en qué consisten cada una de las aplicaciones. La **sección 2** presenta el traductor de Braille, la **sección 3** está dedicada a la aplicación de reconocimiento de formas gráficas y por último la **sección 4** finaliza con las conclusiones.

2. Traductor de Braille a vista

A continuación presentamos el entorno en que se ha desarrollado esta aplicación y algunos de los principios tecnológicos en que se basa.

2.1. Utilidades y entorno de la aplicación

El sistema presentado, a diferencia de la mayoría de los que existen en la literatura, trabaja con escritos realizados a mano mediante una máquina de escribir Perkins, en vez de ser textos realizados por impresora de Braille (ver **figura 1**). Actualmente, un niño invidente, que todavía no ha empezado a utilizar el ordenador para realizar sus escritos, utiliza este tipo de máquina. Si está integrado en una escuela donde los profesores y compañeros no entienden Braille es necesario que alguien traduzca estos documentos, para que le corrijan sus ejercicios o compartílos con sus compañeros de clase. Esta tarea la realizan profesores de soporte que van periódicamente a las escuelas. Una traducción simultánea de estos documentos

permite a los niños compartir sus escritos de una manera más dinámica con su entorno. Sus profesores pueden corregir sus trabajos de una manera más directa y ágil, dándoles las notas al mismo tiempo que al resto de alumnos. Sus compañeros pueden utilizar los apuntes del niño facilitando su integración en el aula, y sus familiares pueden realizar un seguimiento del aprendizaje del niño gracias a que pueden entender sus trabajos escritos. El sistema presenta otra ventaja más genérica, la posibilidad de "fotocopiar" un texto Braille. Dado que fotocopiar un texto Braille implica reimprimirlo en relieve, es necesario interpretar qué hay escrito en el texto original y luego reimprimirlo. Actualmente este proceso sólo puede realizarse reescribiendo el texto, pero con el sistema desarrollado se puede escanear, analizar y guardar en formato electrónico sin necesidad de reescribirlo.

El sistema consta de un ordenador personal conectado a un escáner de sobremesa. Desde la aplicación se permite escanear un documento Braille, reconocer los caracteres Braille presentados y traducirlos a vista. La notación Braille es muy extensa y permite escribir, además de textos, música, química, fórmulas matemáticas, etc. Cada uno de estos dominios tiene sus propios símbolos y reglas de emplazamiento para darle diferentes significados a estos símbolos y sacar más partido a un código basado en seis puntos (existe otra codificación Braille de ocho puntos pero no la trataremos en este artículo).

Por otro lado los documentos analizados están realizados por niños mediante una máquina Perkins, lo que significa que pueden tener imperfecciones tales como puntos "borrados", que consisten en puntos aplastados que no se notan al tacto, pero que, en la mayoría de los casos, se ven en la imagen (ver **figura 2b**)-. También pueden aparecer manchas si es un escrito que se ha utilizado mucho pasando la mano sobre él. En ocasiones, las hojas se sacan de la máquina para releerlas y se reintroducen para continuar escribiendo y esto puede suponer un desajuste en la alineación de los caracteres y en su separación con el resto de líneas. Por

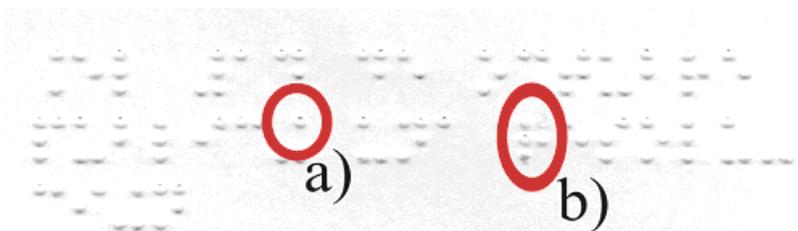


Figura 2. Imperfecciones en los puntos Braille.

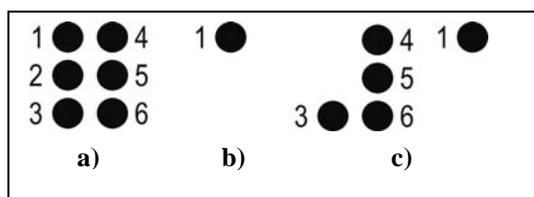


Figura 3. Estructura de los caracteres Braille. a) Símbolo Generador, b) Letra a, c) Número 1.

último la presión ejercida sobre las teclas de la máquina Perkins puede provocar pequeños agujeros en los puntos resaltados, estos agujeros aparecerán en forma de manchas adicionales en la imagen escaneada -ver figura 2a)-. Por todo ello tenemos escritos con diferentes imperfecciones: manchas que pueden confundirse con puntos, puntos borrados que parece que estén marcados, puntos dobles (la sombra del punto y el agujero producido por un marcado demasiado fuerte), desalineación de algunas líneas de texto, etc.

2.2. Tecnología OBR utilizada

Un OCR es un programa de ordenador que permite la conversión automática de imágenes digitales de documentos de texto (adquiridos mediante escáner o cámara digital) a un archivo de texto plausible de ser utilizado por un editor de texto [6] [7]. Actualmente existen muchas herramientas de OCR, a veces distribuidas con escáneres, que permiten digitalizar documentos en papel y traducirlos a un formato legible por el ordenador, como los ficheros ASCII, evitando tener que reescribir esos textos impresos. Desde un punto de vista tecnológico, un sistema de OCR consta de tres etapas: procesamiento de bajo nivel, segmentación/detección de la estructura del documento y reconocimiento de los caracteres o clasificación. El procesamiento de bajo nivel aplica técnicas de procesamiento de imágenes para la mejora de la imagen digital adquirida a partir de papel (aumento del contraste, filtraje para eliminar sombras, manchas y diferentes tipos de ruido que dificultan el reconocimiento, correcciones geométricas, etc.). La segmentación es el proceso de separación de las componentes que conforman la estructura física del documento y que se estructuran de manera jerárquica como: párrafos, líneas, palabras y caracteres. Finalmente, la etapa de reconocimiento o clasificación consiste en identificar, para cada zona de la imagen que contiene un elemento que es un carácter, a qué clase pertenece, es decir, de qué carácter se trata. Este proceso consiste en comparar un conjunto de características extraídas de la imagen [8] con unos patrones o modelos previamente aprendidos en una base de datos de referencia. Desde un punto de vista científico, el reconocimiento óptico de caracteres es un problema superado. Sin embargo todavía quedan algunas líneas abiertas como el reconocimiento omnifluente, es

decir, OCRs capaces de reconocer textos impresos en cualquier tipografía; o el reconocimiento de escritura manuscrita, llamado ICR (*Intelligent Character Recognition*).

El reconocimiento de escritura Braille, *Optical Braille Recognition* (OBR), es un caso particular de OCR. Presenta dos diferencias respecto a los planteamientos clásicos de OCR. En primer lugar el texto no está impreso sino que está formado por relieve que a veces presenta irregularidades en su aspecto. Hemos de tener en cuenta que los textos escritos en Braille no están diseñados para ser vistos sino para ser tocados. Esto implica que el papel utilizado puede presentar marcas visuales que dificulten su análisis mediante visión, como manchas en el papel o caracteres con puntos borrados. En segundo lugar la tipografía no está formada por un conjunto de símbolos como las letras o números, sino que los caracteres se forman a partir de una matriz de seis puntos distribuidos en tres filas y dos columnas o una secuencia de los mismos. Esta matriz es el símbolo generador -ver figura 3a)-. La codificación de cada carácter se establece en base a los puntos que están activos en cada momento y a la secuencia en que se encuentran. Por ejemplo, el símbolo utilizado para designar la letra 'a' y el número '1', -ver figuras 3b) y 3c)- es el mismo, sólo que en el segundo caso tenemos otro carácter previo que nos indica que lo que tenemos a continuación es un número. Según esto, segmentar caracteres Braille en la imagen significa encontrar y agrupar los puntos que forman un carácter, que aparecerán como pequeñas sombras provocadas por su relieve y la iluminación del escáner. Una vez segmentados los caracteres Braille, interpretarlos significa: primero clasificarlos en función de los puntos que tengan activos y después estudiar la secuencia de puntos para poder darles una interpretación.

Existen diferentes artículos en la literatura sobre reconocimiento de escritura Braille [5]. El sistema presentado en este artículo consta de los siguientes pasos.

a) Escaneado de la imagen: La página escrita en Braille se escanea utilizando un escáner estándar de sobremesa. Es recomendable que se escanee a 200 dpi's, para que el análisis posterior de la imagen reciba información suficiente. Asimismo, las páginas deben digitalizarse en posición vertical, aun-

que puede haber una pequeña variación de ± 3 grados debido a su colocación manual. Si la resolución de escaneo es demasiado baja no tendremos suficiente información para poder analizar la imagen, pero si es demasiado alta la imagen será excesivamente grande. Ver figura 4 como ejemplo de imagen escaneada.

b) Filtrado de la imagen: En la imagen del documento escaneado aparecen los puntos resaltados, pero también pequeñas manchas y otras marcas no deseadas. En este proceso se eliminan, en la medida de lo posible, las marcas que no pertenecen a puntos reales. Al final del proceso nos quedamos con una imagen binaria en la que tenemos dos valores correspondientes a: fondo y puntos de los caracteres Braille. Para conseguirlo se utilizan técnicas de filtraje de la imagen: primero una binarización adaptativa de la imagen para clasificar los píxeles en dos clases: parte de punto o fondo. Una vez extraídos los píxeles que pueden formar un punto Braille se agrupan en regiones conexas, utilizando técnicas de morfología matemática, y se calcula el área de cada una. Finalmente se seleccionan aquellas regiones cuyas áreas estén entre unos determinados parámetros. En la figura 5 se muestra un ejemplo.

c) Extracción de las posiciones de los caracteres: Una vez tenemos una imagen con los píxeles que forman los puntos Braille encontramos la alineación de estos puntos y los agrupamos siguiendo el patrón del símbolo generador de la figura 3a). La figura 6 muestra las estructuras de seis posiciones encontradas en un documento ejemplo. En general, los escritos en Braille tienen unas medidas preestablecidas en cuanto a distancia entre puntos de un mismo carácter en horizontal y vertical y distancias entre caracteres en horizontal y vertical. Estas medidas se han seleccionado para que la lectura de los escritos sea lo más cómoda posible. Por ello, si el escaneo de la imagen se realiza siempre a la misma resolución podemos establecer cuales serán estas medidas. Sin embargo, al estar analizando las sombras de los puntos y no los puntos en sí, se producen deformaciones y variaciones en estos parámetros. Además pueden aparecer errores debidos al calibrado de la máquina Perkins y el hecho de quitar y volver a poner una página en la máquina de escribir puede hacer variar la orientación y disposición de unos bloques de líneas con respecto a otros. Por todo ello el análisis no es trivial y pueden producirse errores. Algunas de las técnicas utilizadas son: la transformada de Hough [9] para determinar alineaciones de puntos y el análisis del histograma de mapas de distancias de puntos para determinar automáticamente cuales son las distancias intra e inter caracteres tanto horizontales como verticales.

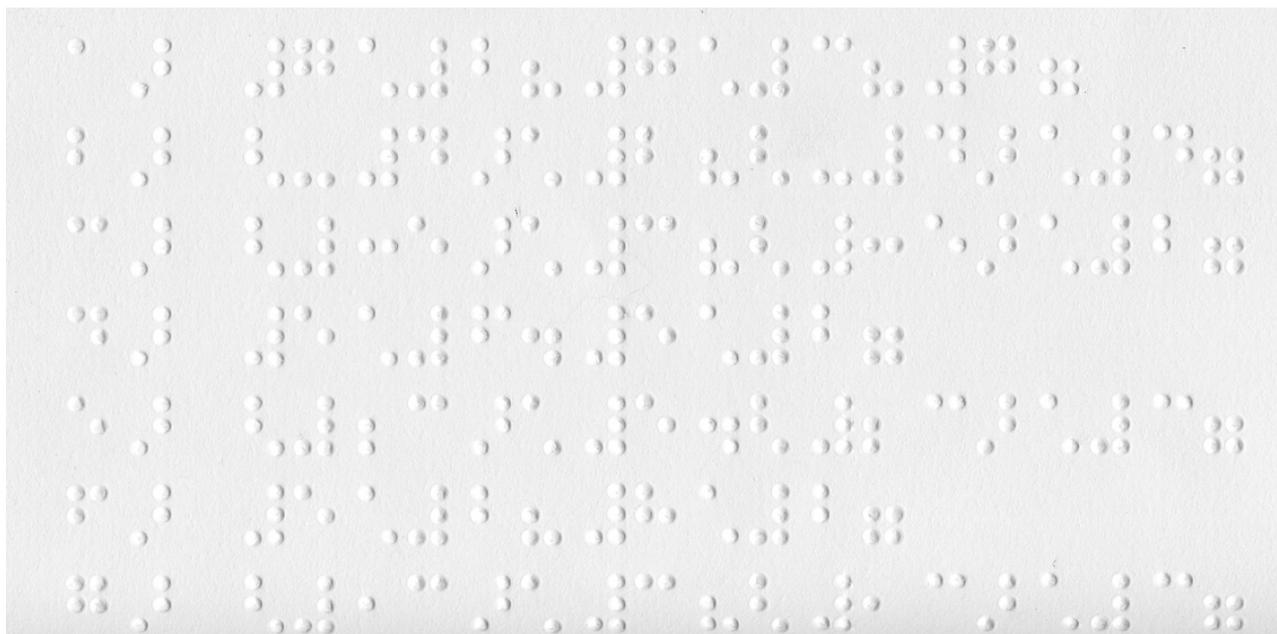


Figura 4. Imagen escaneada de documento Braille.

d) Reconocimiento de los caracteres: Una vez agrupados los puntos en caracteres los codificamos según los puntos que tienen activos en el tramado de seis. Para cada posible distribución de puntos, una tabla nos indica si la codificación es válida y su interpretación, en función de los diferentes modos que podemos encontrar. Inicialmente tenemos los modos: texto o matemáticas, castellano o catalán, minúscula o mayúscula. En futuras versiones podrían añadirse otros idiomas y otros lenguajes como el musical o el de química. La diferenciación entre idiomas es necesaria para los caracteres especiales como la ñ la ç, las vocales acentuadas o con diéresis, etc. Mientras que las matemáticas, química o música tienen símbolos especiales y combinaciones de símbolos con significados propios.

e) Traducción de la secuencia de caracteres: Una vez encontrada la posición de la tabla correspondiente a un carácter Braille encontrado, obtenemos su traducción en función del modo en que nos encontramos. Existen diferentes caracteres que nos indican el modo. Para empezar hemos de presuponer letras minúsculas y un determinado lenguaje, por ejemplo el castellano, a continuación caracteres especiales nos indicaran que pasamos a mayúsculas o cursiva, e incluso que estamos empezando una fórmula matemática. La estructura que debe seguir la disposición de caracteres puede ser representada mediante una gramática que indique qué secuencia de símbolos es correcta y cual no, modificando el modo cada vez que encuentre un símbolo especial de (mayúscula, cursiva, matemáticas, etc.). Siguiendo la gra-

mática, y anotando el modo en que estamos en cada momento, se analizará cada carácter yendo a la posición de la tabla que le corresponda y seleccionando la traducción que le corresponda según el modo en que estemos o cambiando de modo cuando se trate de un símbolo especial de cambio de modo. Si encontramos un error sintáctico en la transcripción, por ejemplo un paréntesis que se ha abierto en una fórmula matemática y no se ha cerrado, se puede dar un error y seguir analizando el resto del documento. Esto es importante al tratarse de documentos que pueden contener errores pero que es necesario dejarlos para que puedan ser corregidos por el profesor.

3. Interpretador de dibujos a mano alzada

En este apartado presentaremos la utilidad de este sistema y sus diferentes aplicaciones, así como una descripción de las principales técnicas utilizadas.

3.1. Utilidad y entorno de la aplicación

Habitualmente los niños invidentes empiezan a dibujar con un lápiz sobre un papel plastificado, de manera que el dibujo se puede ver y tocar su relieve. Por otro lado los ordenadores pueden "ver" los dibujos realizados sobre una tableta digitalizadora. La idea es unir estas dos formas de escritura poniendo un papel plastificado sobre una tableta digitalizadora y dibujando encima con el lápiz de la tableta. De esta forma el niño puede tocar el dibujo realizado y el ordenador tiene las coordenadas de los trazos dibujados. El sistema presentado permite a los profesores crear fichas de trabajo que enseñen a diseñar, de forma interactiva, fi-

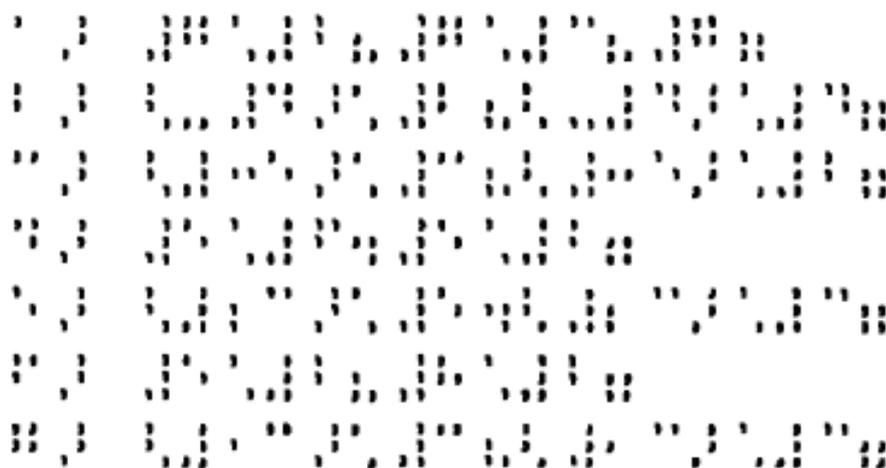


Figura 5. Extracción de puntos de la imagen original que aparece en la figura 4.

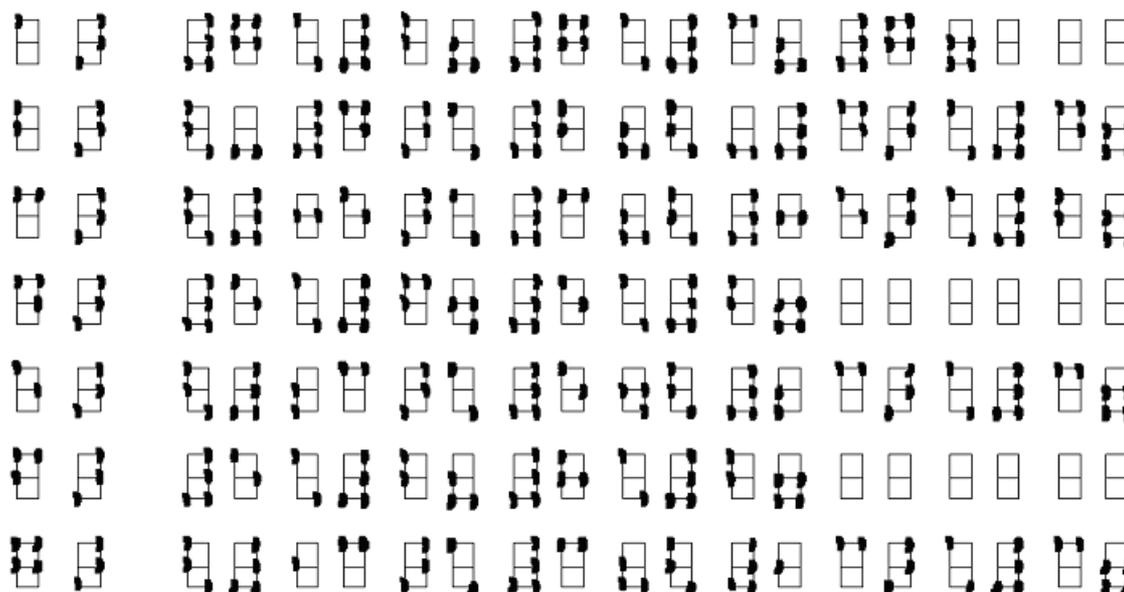


Figura 6. Agrupación en caracteres de los puntos encontrados en la figura 5.

guras geométricas o cualquier símbolo gráfico hecho a partir de líneas.

Además puede utilizarse para incluir en documentos diseños gráficos realizados a mano alzada cuyas posibles imperfecciones hayan sido corregidas. Las fichas consisten en diferentes casillas que tienen asociadas unas funciones. Por ejemplo, una imagen y la opción de reseguirla, además de un mensaje que se oirá cuando el dibujo realizado sea correcto y otro cuando sea incorrecto. Una vez diseñada la ficha, ésta puede imprimirse con una impresora especial que realiza relieves además de colorear los dibujos. La ficha impresa en papel plastificado se coloca sobre la tableta digitalizadora y el niño puede seleccionar táctilmente cada una de las casillas de la ficha y oír un mensaje que le indica la acción que ha de realizar: reseguir la figura que aparece o copiarla en una casilla diferente. El niño reconoce la figura al tocar su relieve en el papel y la dibuja marcando con el lápiz sobre el papel que está sobre la

tableta digitalizadora. Esto hace que el dibujo sea visible para el niño y para el ordenador. A medida que el niño dibuja, el ordenador analiza los trazos y determina si la figura se parece a la esperada o no, interpretándola en base a un modelo conocido como un cuadrado, triángulo, silueta de animal, letra mayúscula, etc. El sistema puede dar una respuesta auditiva y visual que el profesor puede especificar. Los mensajes visuales permiten que el sistema también pueda ser utilizado por niños con resto visual o por el resto de sus compañeros no invidentes.

Ver un esquema en la figura 7. Un estudio interesante sobre interfaces similares puede encontrarse en [10].

3.2. Entorno basado en una interfaz caligráfica

Este sistema se basa en el reconocimiento de símbolos gráficos, que puede definirse como la localización e interpretación de estructu-

ras visuales en documentos gráficos según el dominio en que aparecen. Así, una misma estructura gráfica puede tener interpretaciones diferentes según el contexto en que aparezca. Podemos clasificar el reconocimiento de símbolos gráficos en dos tipos según la manera de introducir la información: reconocimiento *off-line* o estático y reconocimiento *on-line* o dinámico. El primero, parte del documento ya escaneado y sólo tiene información de los trazos que aparecen en la imagen. El segundo, obtiene el documento de forma interactiva, por ejemplo mediante una tableta digitalizadora, un tablet PC o una PDA, obteniendo información no sólo de los trazos de la imagen sino de la secuencia temporal de su creación, la velocidad, etc. El trabajo que presentamos se basa en el reconocimiento *on-line* de símbolos gráficos introducidos a mano alzada mediante una tableta digitalizadora. Este tipo de interfaces se llaman interfaces caligráficas (*sketching interfaces*), y pueden verse otros trabajos

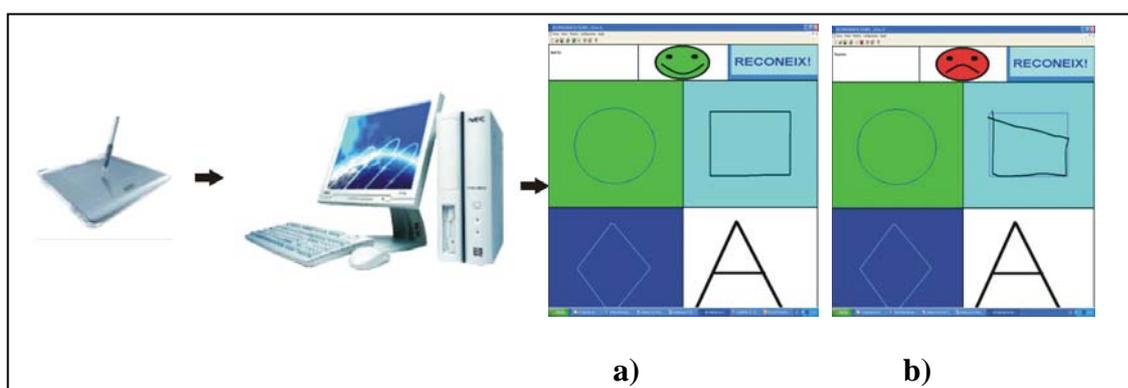


Figura 7. Sistema de reconocimiento gráfico. a) Respuesta correcta, b) respuesta incorrecta.

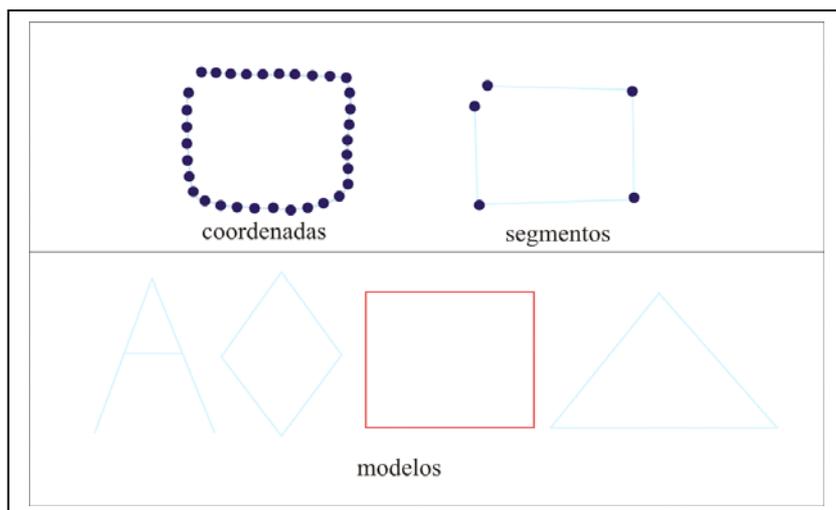


Figura 8. Presentación de las coordenadas introducidas en la tableta digitalizadora, segmentos extraídos y posibles modelos a comparar.

relacionados en [11] [12]. El reconocimiento *on-line* se descompone en tres fases: adquisición, reconocimiento de las primitivas e interpretación. La adquisición consiste en extraer del dispositivo gráfico la información de los trazos y codificarla en forma de secuencias de coordenadas o en las líneas que las aproximan junto con las características que añade el dispositivo de velocidad, etc. El reconocimiento de primitivas consiste en la asociación de trazos a estructuras gráficas simples como polilíneas, arcos de circunferencia, o figuras geométricas simples. Finalmente la interpretación consiste en determinar el significado de la estructura gráfica dibujada y formada por las primitivas de más bajo nivel. Para ello es necesario utilizar la información semántica del entorno en el que se trabaja. A continuación describiremos las diferentes fases del sistema presentado.

a) Adquisición de los trazos de dibujo: A medida que los trazos se van realizando el ordenador recibe las coordenadas por las que va pasando el lápiz de la tableta digitalizadora. De esta manera para cada posible línea del dibujo podemos tener una serie de coordenadas (ver **figura 8**) Una vez leídas las coordenadas debemos analizar qué segmentos están formando, analizando los puntos en los que la curvatura es mayor y en los que la velocidad de los trazos cambió. Una vez analizado esto podemos determinar cuáles son los segmentos dibujados como los presentados en la **figura 8**.

b) Reconocimiento de las primitivas: Los segmentos extraídos se agrupan en primitivas. Estas primitivas dependen del dominio de la aplicación. Por ejemplo si estamos trabajando con diagramas, una primitiva podría ser un círculo, una flecha, un rombo, etc. En nuestro

caso las primitivas son las figuras geométricas que podemos dibujar, por ejemplo, cuadrados, rombos, círculos, segmentos, etc. En el caso mostrado en la **figura 8** la primitiva que hemos encontrado es una polilínea abierta que se parece a un cuadrado.

c) Interpretación del diseño gráfico: Para finalizar se compara el dibujo con uno o varios modelos posibles. En el caso de las casillas que obligan a realizar un dibujo predeterminado, se compara el dibujo realizado con el modelo establecido y se decide si la distorsión encontrada (que siempre existe al ser un dibujo a mano alzada) es tolerable o no. En el caso de dibujos libres se compara la forma dibujada con los modelos almacenados *a priori* y se determina a cual de ellos se parece más y si ese parecido es suficiente para determinar que se trata del mismo dibujo. Los modelos se generan a partir de una secuencia de ejemplos, que sirve para determinar la tolerancia a distorsiones, creando para cada modelo el conjunto de primitivas que lo han de formar y las relaciones topológicas entre ellas. Las comparaciones se realizan detectando qué primitivas contiene el diseño dibujado y las relaciones entre ellas.

4. Conclusiones

Los sistemas presentados son una aplicación de la Visión por Computador al aprendizaje de niños invidentes. El primero aplica técnicas de OBR para traducir un escrito Braille a vista, integrando un módulo de síntesis de voz para comunicarse con el niño. Este sistema favorece la integración del niño en una escuela. El segundo aplica técnicas de reconocimiento de símbolos gráficos en una herramienta que permite el aprendizaje autónomo de figuras gráficas por parte de niños invidentes. Ambos sistemas tienen un

beneficio adicional. El primero permite "fotocopiar" documentos Braille que estén en soporte papel, mientras que el segundo permite retocar diseños gráficos realizados a mano alzada para introducirlos en documentos escritos. Actualmente ambos sistemas están en fase de desarrollo después de la creación de un primer prototipo.

Referencias

[1] T. Sakai. "A history and evolution of document information processing", *Proceedings of 2nd International Conference on Document Analysis and Recognition, volume 1*, pages 377-384, Octubre, 1993.

[2] V.K. Govindan, A.P. Shivaprasad. "Character Recognition - A Review". *Pattern Recognition, Vol. 23, No. 7*, pp. 671-683, 1990.

[3] J. Lladós, E. Valveny, G. Sánchez, E. Martí. "Symbol recognition: current advances and perspectives", *Graphics Recognition: Algorithms and Applications*, Y.B. Kwon, D. Blostein eds. *Lecture Notes in Computer Science, vol 2390*. Springer-Verlag, 2002.

[4] W. Liu. "On-Line Graphics Recognition: State-of-the-Art". En el libro "*Graphics Recognition. Recent advances and Perspectives*", J. Lladós and Y.B. Kwon (eds). LNCS-3088, Springer-Verlag, 2004 ISBN: 978-1-84628-501-1. pp 289-302.

[5] R.T. Ritchings, A. Antonacopoulos, D. Drakopoulos. "Analysis of Scanned Braille Documents". En el libro "*Document Analysis Systems*", A. Dengel, A.L. Spitz (Eds.). World Scientific Publishing Co., 1995, pp. 413-421.

[6] H. Bunke, P.S.P. Wang. "*Handbook of Character Recognition and Document Image Analysis*". World Scientific, 1997. ISBN: 981022270X.

[7] T. Pavlidis, S. Mori. "Optical Character Recognition". *Special Issue of Proceedings of the IEEE, Vol. 80 no. 7*, 1992.

[8] O.D. Trier, A.K. Jain, T. Taxt. "Feature Extraction Methods for Character Recognition - A Survey". *Pattern Recognition, Vol. 29, No. 4*, pp. 641-662, 1996.

[9] P.V.C. Hough. *Method and Means for Recognising Complex Patterns*. Technical Report, 1962. U.S. Patent No. 3069654.

[10] H.M. Kamel and J.A. Landay. "Sketching Images Eyes-free: A Grid-based Dynamic Drawing Tool for The Blind." En *ASSETS 2002: Proceedings of the Fifth International ACM SIGCAPH Conference on Assistive Technologies*, Edinburgh, Scotland, Julio 2002.

[11] J. A. Landay and B. A. Myers. "Sketching Interfaces: Toward More Human Interface Design." *IEEE Computer, vol. 34, no. 3*, Marzo 2001, pp. 56-64.

[12] T. Stahovich, R. Davis, J. Landay. *Proceedings on 2002 AAAI Spring Symposium Sketch Understanding*. Stanford University, Palo Alto. Marzo, 2002.

Notas

1 Este trabajo ha sido financiado por el premio IGC 2006 "Ciudad del Conocimiento" del Ayuntamiento de Barcelona, y por el proyecto de I+D TIN2006-15694-C02-02.