

Desarrollo de un Sistema de Reconocimiento de Marcas basado en Análisis de Contenido en Incertidumbre

Juan F. Inglés-Romero, Cristina Vicente-Chicote

División de Sistemas e Ingeniería Electrónica (DSIE), Universidad Politécnica de Cartagena
Campus Muralla del Mar, s/n. Cartagena 30202,
juanfran.ingles@upct.es, cristina.vicente@upct.es

Resumen

Este artículo describe el desarrollo de un sistema de reconocimiento de marcas dirigido a la lectura automática de documentos de tipo 'test'. El proceso que se propone se basa en el análisis de documentos, previamente digitalizados, haciendo uso de técnicas de procesamiento de imágenes. En primer lugar, se realiza un aprendizaje supervisado del formato del documento y, posteriormente, se analiza su contenido, utilizando vectores de características. El proceso diseñado pretende aportar una solución versátil a las tareas de lectura y valoración de documentos de tipo 'test', que aborde los siguientes retos: (1) robustez ante la variabilidad en el formato del documento, (2) consistencia ante la usual falta de precisión en los mecanismos de fotocopiado y digitalización, y (3) limitación de la incertidumbre asociada al factor humano en el proceso de contestación. Con el fin de demostrar la viabilidad de la propuesta, se ha desarrollado un prototipo en MATLAB [14], cuyos resultados se detallan al final del artículo.

1. Introducción

El uso de cuestionarios de tipo 'test' (en adelante MCQ, *Multiple Choice Questions*) es uno de los medios más utilizados en la actualidad para recabar información en multitud de sectores como la educación, la política, la salud, o el marketing [13], entre muchos otros. En su versión más convencional, los cuestionarios se entregan impresos en papel y son contestados mediante la realización de una serie de marcas, realizadas con lápiz o bolígrafo, sobre zonas predefinidas del cuestionario (casillas de respuesta). Si bien el uso de cuestionarios en formato electrónico empieza a estar cada vez más extendido, la importancia del

papel como medio de captura de datos resulta, a fecha de hoy, incontestable [12].

La traducción de los datos recabados en formato papel a un soporte informático, para su posterior análisis o almacenamiento, supone una tarea muy costosa y propensa a errores si es llevada a cabo de forma manual. Por ello, desde la década de los años 70, se ha trabajado intensamente en el desarrollo y perfeccionamiento de los denominados sistemas OMR (*Optical Mark Recognition*), provistos de un scanner dedicado y que aplican técnicas de reconocimiento de patrones, normalmente basadas en las diferencias de contraste detectadas en posiciones predeterminadas del cuestionario [3].

Una de las principales limitaciones de los sistemas OMR actuales, es su gran dependencia del formato de formulario empleado. Si bien estos sistemas suelen venir acompañados de un software propietario que permite editar ciertos aspectos del formato de los cuestionarios, su flexibilidad es bastante limitada. Así, por ejemplo, suelen requerir el uso de marcas de sincronismo y posicionamiento en los márgenes de los cuestionarios para reducir posibles errores de posicionamiento durante el reconocimiento. Sin embargo, los cuestionarios editados y configurados para una determinada plataforma, no suelen poder exportarse para ser reutilizados en otras. Además, la mayoría de los sistemas OMR suelen presentar una alta sensibilidad al tipo de papel empleado y a los pequeños efectos de desplazamiento, rotación, escalado o variación de contraste del texto que pueden producirse cuando se utiliza un mecanismo de reproducción para imprimir los cuestionarios, ajeno al propio OMR (por ejemplo, una fotocopidora convencional). Por otro lado, los sistemas OMR sólo suelen soportar una única técnica de contestación (por ejemplo, rellenar con lápiz las casillas del

cuestionario). De este modo, si una persona tacha, hace una cruz, o cualquier otro tipo de marca distinta a la esperada en el cuestionario, éste quedaría invalidado y tendría que ser procesado manualmente, reduciéndose sensiblemente la productividad global del sistema. Por último, cabe destacar el elevado coste de la mayoría de los sistemas OMR, siendo esto un grave inconveniente para muchas pequeñas y medianas empresas y organizaciones [2].

El presente trabajo trata de abordar, mediante el uso de técnicas de procesamiento de imágenes, varias de las limitaciones de los sistemas OMR actuales antes descritas. Este artículo se organiza como sigue. En la siguiente sección se delimita el problema, definiendo el formato de los documentos MCQ soportados. En la sección 3 se describen los algoritmos desarrollados; la implementación del sistema y los resultados obtenidos se detallan en la sección 4 y 5; en la sección 6 se revisan los trabajos relacionados; por último, cierra el artículo las conclusiones y los trabajos futuros.

2. Definición preliminar del problema

Dado que no existe un estándar preciso que determine el diseño que deben tener los documentos MCQ, es posible encontrar formatos muy diversos que varían según el tipo de las cuestiones abordadas, el público objetivo, o incluso según el organismo que diseña los formularios. La Tabla 1 recoge algunos de los mecanismos más comúnmente utilizados para marcar las respuestas en distintos tipos de cuestionarios. El objetivo de este trabajo consiste en definir un proceso automático, flexible y robusto para reconocer los distintos tipos de marcas realizadas por los encuestados en los documentos de tipo MCQ. La siguiente sección describe cómo el método propuesto permite abordar cada una de las fuentes de incertidumbre asociadas a la consecución de este objetivo.

2.1. Incertidumbre asociada al formato MCQ

Para limitar el impacto de la gran variedad de formatos existentes y facilitar el diseño del algoritmo de reconocimiento de marcas, se partirá de una serie de premisas, la mayoría de las cuales ya son tenidas en cuenta por los sistemas OMR

actuales. Así, un formulario se considerará válido si está compuesto por:

1. Un cuerpo central formado por una serie de cuestiones, cada una de ellas asociada a un conjunto de dos o más "casillas de respuesta", equiespaciadas y alineadas horizontalmente a su derecha. El encuestado deberá seleccionar y marcar una y sólo una respuesta para cada pregunta. Las casillas de respuesta podrán adoptar diferentes formas, siempre que éstas sean regiones cerradas (por ejemplo, en la Tabla 1, serían válidas las formas *a*, *c* y *e*).
2. Opcionalmente, un encabezado y un pie de página con información gráfica y/o textual, dirigida a poner en contexto y dar instrucciones al encuestado. Para el éxito de los algoritmos de visión diseñados, resulta importante que en los elementos gráficos adicionales presentes en el documento no aparezcan elementos estructurales similares en tamaño y forma a los empleados para las casillas de respuesta.

2.2. Incertidumbre ajena al formato MCQ

Además de la incertidumbre asociada a la gran variedad de formatos de formulario existentes, resulta imprescindible considerar también otras posibles fuentes de incertidumbre, como las asociadas a las alteraciones en su contenido (por ejemplo, causadas durante el proceso de reproducción de los formularios) o a factores humanos.

MUESTRA MÉTODOS DE CONTESTACIÓN																					
a. Rellenado																					
b. Círculos	0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9																				
c. Checkbox	<table border="1"><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr></table>	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9												
0	1	2	3	4	5	6	7	8	9												
d. Ticks	0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9																				
e. Segmentos	<table border="1"><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td><input type="checkbox"/></td><td><input type="checkbox"/></td><td><input type="checkbox"/></td><td><input type="checkbox"/></td><td><input checked="" type="checkbox"/></td></tr></table>	1	2	3	4	5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>										
1	2	3	4	5																	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>																	

Tabla 1. Muestra de distintos mecanismos de contestación de formularios MCQ [11].

En cuanto a la incertidumbre asociada al factor humano, ésta suele estar relacionada con las marcas que realizan los distintos encuestados. Incluso cuando las instrucciones sobre cómo cumplimentar el formulario son claras y precisas, las marcas que realizan distintas personas pueden variar ostensiblemente, incluso las realizadas por una misma persona en distintas preguntas del formulario. El medio de escritura empleado e incluso el mayor o menor interés de la persona por contestar el formulario, determinan en gran medida la precisión de las marcas. Podríamos encontrar, por ejemplo, contestaciones imprecisas que sobresalen de los límites de las casillas de respuesta, dificultando su reconocimiento automático.

En este trabajo consideramos que una persona ha seleccionado una respuesta cuando traza cualquier tipo de marca (cruz, rellenado, tick, etc.) dentro de los límites de una de las casillas de respuesta, quedando fuertemente diferenciada con respecto a las casillas no marcadas. No se restringe el medio de escritura, si bien no se permite el uso de correctores (e.g., typex), ya que su uso puede conllevar la eliminación parcial o total del borde de la casilla de respuesta.

Otra de las posibles fuentes de incertidumbre es la relacionada con los procesos de reproducción y digitalización de los formularios. Así, el uso de fotocopiadoras, impresoras, o escáneres podría alterar el contenido del documento, introduciendo pequeños efectos de rotación, translación, escala, variación del contraste o aparición de ruido impulsional. Supondremos que el contenido del documento puede verse también deformado por estos cambios.

3. Desarrollo del sistema

En esta sección se describen los algoritmos de procesamiento de imágenes diseñados para abordar el problema de reconocimiento de marcas en documentos MCQ, teniendo en cuenta las posibles fuentes de incertidumbre descritas en el apartado anterior.

3.1. Planteamiento adoptado

Partiendo de la imagen digitalizada de un formulario MCQ cumplimentado por uno de los encuestados, y teniendo en cuenta las fuentes de

incertidumbre descritas en la sección anterior, el reto consistiría en (1) detectar la estructura organizativa del documento, esto es, el número de preguntas y posibles respuestas que contiene, y (2) reconocer las marcas realizadas por el encuestado con el fin de inferir cuáles han sido sus repuestas a las preguntas planteadas en el cuestionario. Las fases que proponemos para llevar a cabo este proceso son las siguientes:

1. Una *fase de adquisición* en la que el documento es digitalizado y preprocesado. Esta fase tiene como resultado una imagen binaria.
2. Una *fase de entrenamiento* supervisado para reconocer la estructura organizativa del documento (número de preguntas y número de opciones de respuesta asignadas a cada pregunta), así como para caracterizar las propiedades visuales que definen una casilla de respuesta “no marcada”.
3. Una *fase de procesamiento* en la que se detectan las casillas de respuesta “no marcadas” en el documento y, a partir de ellas, se deduce las opciones que han sido seleccionadas por un encuestado.

La división del proceso en estas etapas permite obtener los siguientes beneficios:

- La fase de entrenamiento permite que el proceso se amolde al formato específico del documento, de forma que el sistema adquiere una gran flexibilidad ante la diversidad de formatos existentes.
- Seleccionar como elemento de interés las casillas de respuesta “no marcadas” nos permite, en cierto modo, aislar la resolución del problema de la incertidumbre asociada al factor humano.
- La robustez del proceso de detección de las opciones no marcadas ante posibles deformaciones introducidas al reproducir y/o digitalizar los formularios, se consigue realizando una buena selección de los componentes del vector de características. Así, por ejemplo, si las casillas de respuesta tienen forma cuadrada, su área permanecerá invariante ante rotaciones y translaciones del documento, pero no ante posibles escalados. En este caso, la elección como propiedad discriminante de los momentos invariantes de Hu o los descriptores de Fourier resultaría más adecuada.

3.2. Entrenamiento de documentos

La fase de entrenamiento permite al sistema abordar el reconocimiento de un nuevo documento MCQ, esta fase es innecesaria si el documento ya ha sido entrenado con éxito previamente. Partiendo de la imagen binaria del formulario no contestado, el objetivo radicaría, por un lado, en caracterizar las casillas de respuesta (en adelante referidas como regiones de interés, *ROI*) y tras ello, analizar la estructura organizativa del documento, identificando tanto las preguntas como las casillas de respuesta.

Respecto a la caracterización de la ROI, dado que este elemento mantiene las peculiaridades indicadas en la sección 2 (esto es, región uniforme delimitada por un borde), su caracterización puede realizarse mediante un vector de características seleccionadas para permitir su discriminación frente al resto de los elementos visuales en el documento. En este trabajo se ha seleccionado un conjunto de características básicas centradas en la forma y el tamaño de los objetos [6], en concreto: el área, el perímetro, la excentricidad, los ejes de la elipse equivalente, los momentos de Hu y los descriptores de Fourier. La correcta combinación de estas características confieren cierta robustez frente a rotación, translación, escalado y cambios de contraste del contenido de los documentos.

Tras la caracterización de la ROI se realiza una búsqueda de los elementos visuales similares presentes en la imagen. Para ello, previamente, se aplica un proceso de etiquetado de los objetos de la imagen con el fin de calcular el vector de características para cada uno de ellos. Los objetos cuya distancia al vector de referencia de la ROI sea menor que un umbral, serán considerados como casillas de respuesta. La expresión de distancia empleada es la siguiente, donde $m(i)$ es el valor de la componente i -ésima del objeto considerado y $n(i)$ la correspondiente de la ROI.

$$e = \frac{1}{M} \sum_{i=1}^M \left| \frac{m(i) - n(i)}{n(i)} \right|$$

Respecto al análisis de la estructura del documento, para inferir el número de opciones de cada pregunta: (1) se parte de las coordenadas del punto central (centroide) de las regiones detectadas como casillas de respuesta y, suponiendo que la orientación del formulario es la correcta (de lo contrario debería aplicarse

previamente el algoritmo de corrección de orientación detallado más adelante) (2) se analiza la distancia relativa entre los centroides, agrupándolos por líneas, cada una de las cuales se corresponderá con una de las preguntas del formulario. De esta forma, resulta sencillo inferir la estructura organizativa del documento MCQ.

3.3. Procesamiento de documentos

Para realizar el reconocimiento de las respuestas marcadas por un encuestado, partimos de las imágenes binarias de los formularios contestados y de la información que resultó en la fase de entrenamiento previa, esto es, el vector de características de la ROI y la estructura organizativa del documento. Los pasos que se seguirían para realizar el procesamiento se describen a continuación.

1. Detección de las casillas de respuesta no marcadas presentes en el documento. Tal y como procedimos durante la fase de entrenamiento, los objetos cuya distancia al vector de referencia de la ROI sea menor que un umbral, serán considerados como casillas de respuesta. De esta manera, se detectarán las casillas de respuesta vacías, dado que al no sufrir modificación por el trazado de una marca mantendría el grado similitud con la ROI.
2. Detección y ajuste de la orientación del contenido del documento (si es necesario).
3. Para cada pregunta del formulario inferimos la respuesta marcada a partir de la posición relativa de los centroides de las casillas vacías detectadas en el paso 1. La Figura 1 representa el diagrama de flujo del algoritmo asociado.

3.4. Corrección de la orientación

Detectados los centroides de las casillas de respuesta no marcadas, su posición en el eje y permitiría inferir líneas, y por lo tanto, diferenciar las opciones asociadas a cada pregunta. Para mantener esta relación espacial se requiere de una orientación correcta del documento. El proceso desarrollado para llevar a cabo la corrección de la orientación se basa en un proceso iterativo de búsqueda del ángulo, en el intervalo $[0, \pi)$, que minimiza la suma de las diferencias de la proyección de los centroides sobre la recta

definida por el ángulo considerado. Matemáticamente, dado el conjunto de centroides C , donde el centroide k -ésimo queda definido por sus coordenadas en el espacio (x_k, y_k) . Para $\varphi \in [0, \pi)$ la orientación del documento será:

$$\varphi_{orientación} = \operatorname{argmin}(e_{\varphi})$$

$$e_{\varphi} = \sum_j p_{\varphi}(j) - p_{\varphi}(j+1)$$

donde $p_{\varphi}(j) = x_j \cdot \cos\varphi - y_j \cdot \operatorname{sen}\varphi$

Para obtener el sentido del documento, es decir, si está o no girado 180 grados, consideramos la premisa, definida en la sección 2, que indica que las casillas de respuesta se sitúan a la derecha de las preguntas. Teniendo en cuenta el número de centroides que cae en cada cuadrante del documento (en el nuevo espacio determinado por el φ resultante), si el cuadrante que contiene el mayor número de centroides es el 2º ó 3º entonces el documento se encontrará girado 180º.

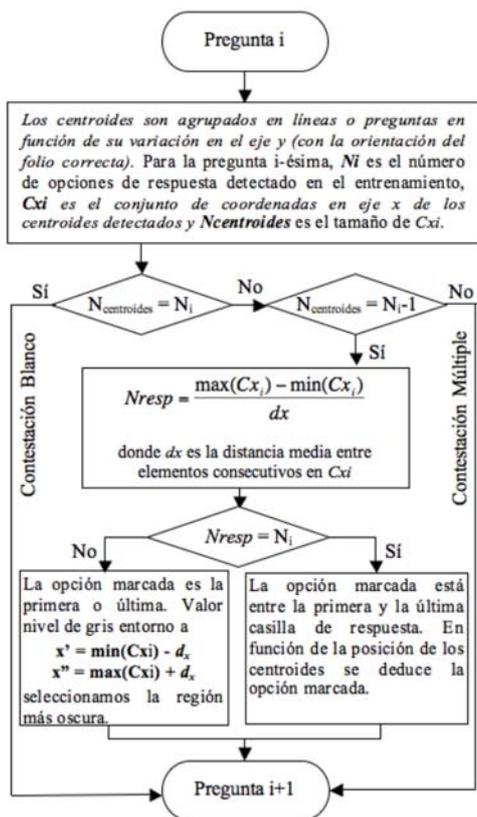


Figura 1. Algoritmo para el reconocimiento de marcas

4. Prototipo implementado

Con el objetivo de demostrar la viabilidad de los algoritmos diseñados, así como la posibilidad de realizar una evaluación con documentos reales. Se ha implementado un prototipo utilizando el entorno de programación MATLAB.

Las imágenes de los documentos a procesar serían introducidas en la aplicación como: (1) ficheros, (2) secuencias de imágenes en formato TIF/TIFF [1], (3) una ruta a un repositorio de imágenes, (4) o desde una cámara. Comentar que excepto la última entrada, la digitalización de los documentos se supone transparente al sistema, ésta puede llevarse a cabo utilizando un escáner cualquiera o un equipo profesional de alta velocidad. Respecto a esta última entrada, se ha empleado una cámara Web común, de resolución 640x480 píxeles, que nos permitirá experimentar en condiciones de gran dificultad con alto grado de incertidumbre en cuanto a ruido y deformación del contenido. La salida de la aplicación consistiría en un fichero de texto tabulado con las frecuencias estadísticas que reflejan el número de veces que una respuesta ha sido marcada por una persona.

En la figura 2 podemos observar el aspecto que tiene la interfaz de usuario del prototipo desarrollado. Los elementos que conforman la aplicación son los siguientes: (1) en el área superior izquierda, un panel con los controles para cargar las imágenes, entrenar y procesar documentos, bajo éste (2) un panel secundario con tres vistas, una primera con la representación en árbol de la estructura organizativa de preguntas y respuestas del documento, una vista de resultados y una vista de configuración. Por último, el mayor elemento de la interfaz lo compondría (3) el visor de imágenes.

4.1. Configuración del entrenamiento

El entrenamiento constituye una fase fundamental para el correcto funcionamiento de los algoritmos de reconocimiento. En la mayoría de los casos, sólo se requerirá una única ejecución de esta fase para obtener resultados satisfactorios, sin embargo, puede que necesitemos repetir el proceso de forma iterativa para ajustar los valores de los parámetros de configuración. Seguidamente se exponen los parámetros más destacados.

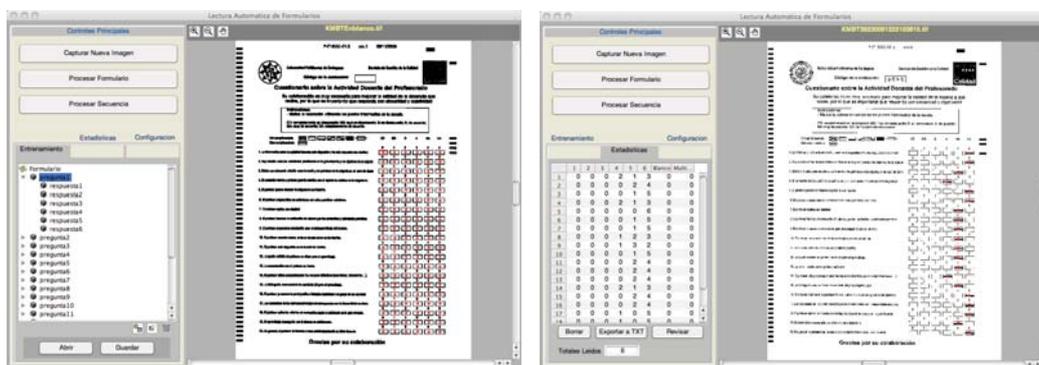


Figura 2. Aspecto de la interfaz de la aplicación. Izda.: vista entrenamiento. Dcha.: procesamiento de un documento

- Tamaño del elemento estructurante (cuadrado) de una operación morfológica de erosión. Esta operación se aplica de forma opcional para preprocesar la imagen de entrada, con el fin de resaltar bordes y hacer más perceptibles las casillas de respuesta en el documento.
- Selección de los componentes que formarán el vector de características a considerar.
- Error máximo empleado para decidir qué objetos son similares a la ROI.
- Activación del proceso de la detección y ajuste de la orientación del documento.

El entrenamiento resultará fallido si tras su ejecución no se detectan todas las casillas de respuesta y/o la agrupación inferida de preguntas y respuestas es incorrecta para el documento en cuestión. En la Figura 2 (izquierda) podemos observar los resultados tras un proceso de entrenamiento, vemos que dicho proceso se ejecutó correctamente dado que se habrían localizado todas las casillas de respuesta (señaladas con asteriscos) y se obtendría una estructura organizativa también correcta (representación en árbol).

5. Resultados experimentales

En esta sección expondremos algunos resultados obtenidos. En primer lugar, se consideran encuestas reales utilizadas en la evaluación de la calidad docente de la Universidad Politécnica de Cartagena. Tras ello, abordamos el

reconocimiento en entornos complejos usando la cámara como medio de captura.

5.1. Reconocimiento de encuestas reales

La Figura 3 (izquierda) muestra el formato típico de los documentos MCQ considerados. Todos ellos contienen 20 preguntas con un número de respuestas/pregunta cambiante. El proceso de digitalización de estos documentos se ha llevado a cabo por medio de un escáner comercial de alta velocidad. De manera automática, el escáner codifica un lote de encuestas como una secuencia de imágenes en un fichero TIF. Las imágenes son monocromáticas de resolución 2333x1653 píxeles.

Los parámetros configurados han sido los siguientes: ajuste de orientación desactivada, operación de erosión con elemento estructurante cuadrado de 4 píxeles, umbral de error de 0.15 y vector de características compuesto únicamente por el "área". Respecto a éste, se han probado diferentes configuraciones de características obteniéndose resultados muy similares.

Tras realizar una única iteración de entrenamiento, se ha procesado un primer conjunto de 122 instancias contestadas por diferentes personas usando rotulador como medio de escritura, el criterio de contestación se basó en trazado de "segmentos" (ver Tabla 1). Para este conjunto, la probabilidad de que todas las marcas en una misma hoja sean reconocidas correctamente resultó del 97,54%. Se ha procesado un segundo conjunto de 20 instancias contestadas empleando diferentes medios de

escritura (lápiz, bolígrafo y rotulador) y un criterio de marcado arbitrario. En este caso la probabilidad de éxito descendió al 80%.

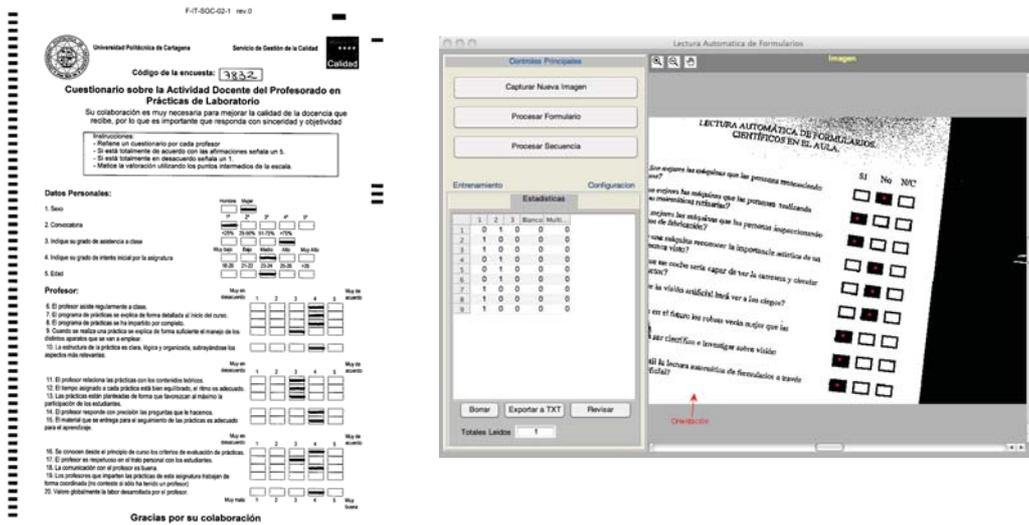


Figura 3. Izda.: ejemplo documento MCQ. Dcha.: vista procesamiento con cámara.

5.2. Reconocimiento a través de la cámara

El documento MCQ empleado en este caso fue diseñado especialmente con un tipo de letra algo mayor de lo normal (14). Se trata de un formulario de nueve preguntas y tres opciones de respuesta cada una. Las características configuradas como discriminantes fueron la relación de ejes de la elipse equivalente y el área, además se activó el ajuste de la orientación del documento.

La fase de entrenamiento resultó más compleja, puesto que las imágenes captadas eran más susceptibles a introducir fallos dada la variabilidad del entorno y el procedimiento de exposición de los documentos ante la cámara. Así, por ejemplo, la iluminación fue determinante, por otro lado, también pueden introducirse en la imagen objetos ajenos al documento, presentes en el fondo de la escena, que ocasionarían fallos si éstos adoptan características visuales similares a la ROI. Por todo ello, normalmente, requerimos realizar varias iteraciones de entrenamiento para conseguir los ajustes más adecuados al entorno.

En la Figura 3 (derecha) se muestra la aplicación tras la fase de procesamiento. En cuanto a esta fase, aparte de los efectos adversos que, al igual que el entrenamiento, afectaron (e.g. la iluminación), debemos añadir la variabilidad de la distancia de exposición de los documentos a la

cámara. Ésta repercute como aspecto importante, sobre todo si consideramos que la sujeción del papel fue manual, por ello, podría aparecer un efecto de escalado en el contenido del documento (según se acerque más o menos el papel a la cámara), lo que hace recomendable seleccionar, en este caso, medidas robustas a la escala. Por todos estos efectos, las pruebas realizadas presentaron resultados pobres y muy influenciados por las condiciones de captura y del entorno.

6. Trabajos relacionados

Trabajos relacionados con el uso de sistemas OMR aplicados a problemas particulares son, Ferrari G. et al. [4], dirigido al almacenamiento de información epidemiológica, [8] para la gestión de información de casos médicos de emergencias o [9] para la corrección automática de exámenes y lectura de encuestas en el entorno universitario. Estos sistemas adolecerían de los problemas presentados en la introducción, principalmente la escasa flexibilidad ante diferentes formatos y la falta de robustez ante variaciones en el contenido. Por otro lado, los sistemas OMR presentes en el mercado (véase [2,9]), tienen como desventaja, frente al método propuesto en este trabajo, sus altos precios, que puede deberse al hardware especializado que emplean. En este sentido, existen trabajos que proponen alternativas de bajo

coste basadas en procesamiento de imagen, a través de una solución software y un scanner común [10], o una solución hardware basada en FPGAs y un sensor CCD de alta resolución [7]. Sin embargo, de nuevo, ninguno de estos trabajos plantean una solución flexible y robusta, así por ejemplo, al igual que los sistemas OMR típicos, proponen soluciones que siguen utilizando marcas de referencia fijas situadas en los márgenes de los documentos. También podemos encontrar otros sistemas comerciales [5], basados en procesamiento de imágenes, que adoptan una perspectiva más general, no obstante, se ha comprobado que el entrenamiento de nuevos formatos puede resultar algo pesado y tedioso en relación al planteado en este trabajo.

7. Conclusiones y trabajos futuros

En una sociedad asaltada por las Tecnologías de la Información y Comunicación, los documentos de tipo MCQ, en su versión más convencional, impresos en papel, siguen dominando claramente las actividades de recolección de información en multitud de áreas. En este trabajo se ha propuesto una solución versátil, robusta y de bajo coste para el reconocimiento automático de documentos tipo MCQ utilizando procesamiento de imágenes. Para ello, se ha diseñado un algoritmo de aprendizaje y de procesamiento de documentos basados en la detección de casillas de respuesta no marcadas, permitiendo así, limitar efectos de incertidumbre.

Las pruebas empíricas reflejan un buen comportamiento de los algoritmos diseñados ante rotaciones, translaciones, escala y cambios de contraste del contenido, así como, una cierta independencia ante la forma de las casillas de respuesta (circulares,...) y los métodos de contestación empleados. Según se ha analizado, la mayoría de los casos erróneos hallados son provocados por defectos en los bordes de las casillas de respuesta, fragmentaciones que hacen que se desvirtúe el objeto de interés. Además, se han contrastado las posibilidades de la aplicación ante entornos no estructurados de gran incertidumbre usando una cámara, obteniéndose una buena adaptación ante deformaciones del contenido y cambios de orientación.

Como trabajos futuros proponemos el desarrollo de un proceso de selección automática de características discriminantes que facilite el

entrenamiento ante nuevos documentos, así como el incremento de la consistencia ante fragmentaciones de los bordes que delimitan la región predefinida de contestación. Por último, dada la utilidad de la herramienta cabría plantearse la implementación de un prototipo en un lenguaje compilado como C/C++, que permita, además, la evaluación del rendimiento de los algoritmos planteados.

Referencias

- [1] Adobe TIFF, <http://partners.adobe.com/public/developer/tiff>
- [2] Axiome Alpha SA. OMR: optical mark reader: www.axiome.ch/OMR/OMRHARDWARE
- [3] Bergeron B.P., *Optical mark recognition*. Postgraduate Medicine online. 2006, 104(2).
- [4] Ferrari G, Zywiets C, Willems JL, Fayn J, Malossi C, Poeta M, Assanelli D. *User interface database for digital SCP-ECG and epidemiological information*. Computers in cardiology. Los Alamitos, CA, USA: IEEE Computer Society Press; 1995. p. 177–80.
- [5] Gravic, <http://www.gravic.com/>
- [6] Gonzalez R.C., Woods R.E., *Digital image processing*. 2nd ed. Prentice Hall, 2002.
- [7] Hussmann S., Deng P.W., *A high-speed optical mark reader hardware implementation at lowcost using programmable logic*, Real Time Imaging, vol. 11, 2005, p. 19-30.
- [8] Inglis R., Windolf J., Pannike A. *Can automatic 'scoring' improve comparability of cases and statistical results in emergency medicine?* Medical informatics Europe. Berlin, Germany: Springer; 1991. p. 973–77.
- [9] NCS Pearson Inc., *OpScans optical mark read (OMR) scanners*. Web: www.pearsonnncs.com
- [10] Saengtongsrikamon C., Meesad S. *Scanner-Based Optical Mark Recognition*. Information Technology Journal, vol. 9.
- [11] Sellen, A.J., Harper. *The Myth of paperless office*. 2001, MIT Press.
- [12] Singh, G., Findlater, L., Toyama K., Helmer S., Rikin G., *Numeric Paper Forms for NGOs*. ICTD 2009. p. 406-416.
- [13] Smith A.M., *Optical mark reading—making it easy for users*. 9th ACM SIGUCCS conference on user services, Atlanta, Georgia, USA, 1981. ISBN:0-89791-054-0, p. 257–63.
- [14] The MathWorks, Inc.: www.mathworks.com