

Medición de la altura del cuerpo medio en la escritura

Carlos F. Romero, Carlos M. Travieso, Jesús B. Alonso, Miguel A. Ferrer.
Dpto. de Señales y Comunicaciones, Universidad de Las Palmas de Gran Canaria
Campus de Tafira, 35017 Las Palmas de Gran Canaria, España.
Tel.: +34 928 452 864 Fax: +34 928 451 243.
email: fabian_romero@ciudad.com.ar; {ctravieso, jalonso, mferrer}@dsc.ulpgc.es

Resumen— En este trabajo se presenta un método efectivo para medir la altura promedio del cuerpo medio de la escritura, analizando las palabras segmentadas desde muestras manuscritas. A través del procesamiento digital de imágenes se calcula la distancia existente entre las líneas de base superior e inferior de la escritura, las cuales son aproximadas utilizando el método de los mínimos cuadrados; dando como resultado el ajuste de dichas líneas rectas por medio de las coordenadas (x,y) de los máximos y mínimos detectados en el contorno de la palabra.

En este experimento, fueron utilizadas 10 muestras de escrituras de 5 escritores, obteniendo una tasa de mediciones correctas máxima del 78%.

Abstract— In this work, an effective method is presented to measure the average height from the middle body of the script, analyzing the words segmented from handwriting samples. The existent distance among the upper and lower baselines is estimated through the digital image processing, which are approximate using the method of the least square; giving the adjustment of this straight lines as a result by means the coordinates (x,y) of the maximum and minimum detection in the contour of the word.

In this experiment, 10 samples of 5 writer's writings were used, obtaining a rate of maxim correct measurement of 78%.

I. INTRODUCCIÓN

Pese a la proliferación de los documentos electrónicos, es incuestionable la gran importancia que sigue teniendo en nuestros días el documento escrito y la propia escritura en sí, haciendo que la identificación y la verificación de escritor tome un papel importante.

La investigación se centraliza en la identificación *off-line* de escritores de documentos manuscritos, por lo cual en el trabajo se propone la medición de la altura del cuerpo medio de escritura de las palabras segmentadas del texto manuscrito, para posteriormente utilizarla como característica del escritor en la etapa de reconocimiento.

Según la metodología desarrollada por la grafología [1], la forma más rápida y sencilla de realizar la medición, es trazando dos líneas que se apoyen con la mayor aproximación posible sobre las zonas superior e inferior de las letras, siendo

la distancia entre dichas líneas, un valor bastante certero de la altura del cuerpo medio de las palabras. Es necesario resaltar que normalmente la línea inferior es más fácilmente detectable que la línea superior. Este hecho se debe a que los escritores suelen utilizar una línea inferior de referencia cuando escriben (real o imaginaria), mientras que la línea superior no existe, y no se asocia a una línea de referencia sino a un proceso derivado de la uniformidad del texto. De esta manera, cuanto más uniformemente escriba un escritor más clara será la línea superior.

Nuestra línea de investigación apunta al desarrollo de una herramienta software para ayudar a los peritos caligráficos en el análisis de textos manuscritos. Esto se realiza mediante la comparación de un documento cuestionado con una muestra de escritura conocida, perteneciente a una base de datos; lo que requerirá la utilización de técnicas de procesamiento de imagen y reconocimiento de patrones [2-4].

Todos los métodos desarrollados en los trabajos publicados en este campo [5-8], dependen de tres etapas básicas:

- Preprocesamiento: eliminación del ruido en la imagen y segmentación del texto, extrayendo las líneas de escritura y posteriormente las palabras.
- Extracción de características: son medidas cuantitativas y cualitativas que permiten diferenciar a los escritores entre sí; pueden ser globales o locales y estructurales o estadísticas.
- Clasificación/reconocimiento de patrones: se busca al escritor más cercano, por medio de las características extraídas.

El resto del artículo está organizado de la siguiente manera, en la sección 2, se describe brevemente la creación de la base de datos. En la siguiente sección se explica cómo son preprocesadas las muestras y cómo se segmenta el texto en líneas de escrituras y éstas a su vez en palabras. En la sección 4 se realiza una descripción del método propuesto, y finalmente, en la sección 5 se redacta una conclusión del trabajo realizado.

II. CREACIÓN DE LA BASE DE DATOS

Se utilizó un texto de 15 líneas para que queden reflejados los hábitos normales de escritura de cada persona.

* Revista Argentina de Trabajos Estudiantiles. Patrocinada por la IEEE.

Para tomar las muestras, se les pidió a los 5 escritores que copiaran 10 veces el texto en su escritura más natural, usando hojas blancas A4 sin líneas y un bolígrafo negro. Las 10 repeticiones permiten determinar la variación de escritura propia del escritor que será necesaria cuando se aborde la etapa de clasificación. Las condiciones en la toma de las muestras fueron normalizadas: la misma clase de papel, bolígrafo y tipo de apoyo.

A continuación, cada una de las muestras fueron escaneadas en una resolución de 300 dpi, obteniendo imágenes en escala de grises.

III. PREPROCESADO Y SEGMENTACIÓN

La finalidad del preprocesamiento es preparar y modificar las imágenes para que el módulo de segmentación produzca los resultados deseados. La segmentación separa las zonas de interés (palabras en nuestro caso) y es un paso clave para el éxito o fracaso del análisis de las imágenes, ver Fig. 1.

La primera etapa del preprocesado consiste en utilizar el método de Otsu, que nos permite determinar el valor gris umbral necesario para realizar la binarización de las muestras [9].

Como consecuencia de la binarización, en la mayoría de los casos el trazo de escritura queda pixelado, es decir poco consistente en algunas partes; por este motivo se realiza otro preprocesado que permite suavizar el trazo para que quede bien definido y también elimina el ruido existente en las imágenes de las muestras.

Se utiliza la transformada de Hough para obtener las coordenadas que permitirán realizar la segmentación del texto en sus líneas de escrituras [10-11]; y a cada subimagen obtenida se le corrige la inclinación (denominada Skew en el ámbito científico), realizando la proyección horizontal generalizada [12] a las imágenes resultantes de rotar dicha subimagen entre -5 y 5 grados; y finalmente buscando la imagen que presente la proyección generalizada con el pico máximo de mayor valor, porque ello implica el menor Skew de la línea de escritura con respecto a la horizontal.



Fig. 1. Obtención de las palabras.

Como paso previo a la separación de las palabras o componentes conectadas, se realiza la detección y eliminación de los signos de puntuación (puntos, acentos y comas).

Finalmente, se deberá segmentar las palabras que componen las líneas de escrituras y para ello hay que establecer los límites de cada una de las palabras; para esta estimación se utilizó el método de las cajas acotadas [13-14] que nos proporciona las coordenadas que permitirán segmentar las palabras. Las cajas acotadas se definen como el mínimo rectángulo que contiene a la componente conectada. A cada palabra segmentada se le aplica nuevamente una corrección del Skew.

En un principio se pensó medir la altura del cuerpo medio a las líneas de escritura, pero se observó que las líneas no son siempre uniformes, es decir, además de ascendentes o descendentes también pueden ser cóncavas, convexas, onduladas o incluso encontrarse el caso de palabras ascendentes y descendentes en una misma línea; por lo que se debería realizar algún preprocesado para alinear dichas irregularidades de las líneas de escritura, añadiendo mayor dificultad a la medición que se pretende hacer.

IV. MÉTODO PROPUESTO

Como se expuso anteriormente se pretende extraer como característica la altura del cuerpo medio de las palabras. La idea es estimar la línea base superior e inferior por medio de los máximos y mínimos, y medir la distancia entre ellas, ver Fig. 2. En una primera prueba se trató de utilizar la proyección generalizada horizontal, a través de la cual, determinando en el histograma un umbral, se ubica la posición de las líneas base. Pero debido a la problemática de que los escritores tienen estilos de escritura diferentes, dicho umbral funciona bien para algunas escrituras y con errores considerables para otras.

Para la obtención de las coordenadas de los máximos y mínimos del contorno, en primer lugar se debe tener en cuenta que al desarrollar el trazo, se levanta la punta del bolígrafo produciendo algunas desuniones entre las letras de las palabra, por lo que se realizó una operación morfológica para lograr que el trazo de escritura sea continuo a lo largo de la palabra.

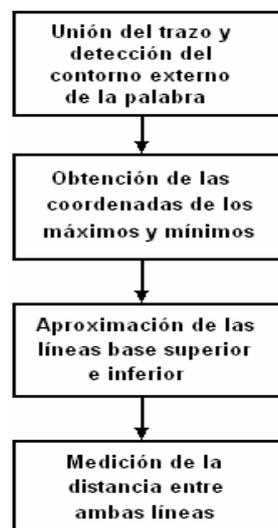


Fig. 2. Etapas de la medición.

El método consiste en posicionarse en el punto de inicio, barrer los píxeles vecinos en el sentido contrario a las agujas del reloj desde la prioridad uno hasta la ocho en búsqueda del píxel vecino que tenga valor uno, al detectarlo lo que se hace es guardar en un vector las coordenadas (x,y) del píxel actual P y posicionarse en el píxel detectado.

Una vez definido el píxel al que se ha saltado, éste pasará a ser el nuevo punto central P, y así este proceso se irá repitiendo hasta retornar al punto de inicio.

Finalmente, se toma el vector donde se han guardado todas las coordenadas de los píxeles del contorno de la palabra y se filtran todos los puntos que estén fuera de las dos bandas (una para filtrar los máximos y otra para los mínimos) cuyos límites son fijados de la siguiente manera:

- Límites de la banda para filtrar máximos: como primer paso se deberá comprobar si la palabra tiene ascendentes, y en el caso de que existan se define como límite superior de la banda a la fila resultante de restar a la componente C_y del centro de masa, el 66% de la distancia entre C_y y el valor mínimo en “y” del vector donde se han guardados las coordenadas (x,y) de los píxeles del contorno; en el caso de que no exista ascendente, el límite será el valor mínimo en “y” de dicho vector.

El límite inferior de la banda se define como la fila resultante de restar a C_y , el 25% de la distancia entre C_y y el límite superior calculado previamente.

- Límites de la banda para filtrar mínimos: igual que en el caso anterior se debe comprobar si la palabra tiene descendentes. Cuando la palabra tiene descendentes se define como límite superior de la banda a la fila resultante de restar a la componente C_y del centro de masa, el 65% de la distancia entre dicha componente y el valor máximo en “y” del vector de coordenadas (x,y) del contorno; sino se toma como límite el valor máximo en “y” del vector.

El límite inferior se define como la fila resultante de sumar C_y y el 35% de la distancia entre el límite superior y dicha componente.

Para comprobar si la palabra tiene ascendentes o descendente se procedió a comparar la componente en “y” del píxel del trazo más alejado de la componente C_y del centro de masa; si dicha proporción es mayor a 0,7 implica que la palabra tiene ascendentes o descendentes.

Todos estos valores umbrales fueron ajustados empíricamente para que se adapten lo mejor posible a los 5 estilos de escritura que componen la base de datos.

Una vez obtenidas las dos bandas se realiza un filtrado guardando en un vector, únicamente las coordenadas de los píxeles del contorno que están dentro de los límites de las bandas determinadas; luego se toman los dos vectores resultantes (Vb_1 y Vb_2) y se les calcula la derivada con

respecto al eje de coordenadas “y”, de la función discreta representada por los valores de cada vector.

$$\frac{\partial Vb_1}{\partial y} = 0 \Rightarrow \text{Máximos} \quad (2)$$

$$\frac{\partial Vb_2}{\partial y} = 0 \Rightarrow \text{Mínimos} \quad (3)$$

Los píxeles donde la derivada sea cero serán los máximos y mínimos buscados.

Para aproximar las líneas base de cada palabra, se decidió utilizar el ajuste de mínimos cuadrados que se basa en hallar la ecuación de la recta que mejor se ajuste a un conjunto “n” de puntos [18-19]. Dicha ecuación es:

$$y = ax + b \quad (4)$$

Donde los coeficientes a y b se determinan mediante regresión lineal con las siguientes fórmulas:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (5)$$

$$b = \frac{\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i}{n} \quad (6)$$

Los valores de x, y son las coordenadas de los mínimos o máximos detectados en el contorno de la palabra, sin tener en cuenta los correspondientes a los ascendentes y descendentes de la palabra porque producirían errores en la estimación correcta de las líneas bases; los mínimos son para aproximar la línea base inferior y los máximos para la superior.

Debido a que las líneas base resultantes en algunos casos pueden tener diferente valor de inclinación, es decir, no son siempre paralelas; se realizan tres mediciones de distancia entre las líneas base (una en el extremo derecho, otra en el izquierdo y la última en el centro de la palabra) y en base a los tres valores resultantes, se calcula la distancia media que será igual a la altura media del cuerpo de escritura de la palabra.

V. CONCLUSIONES

En este artículo se ha presentado la medición del cuerpo medio de las palabras como característica extraída de la escritura de las personas, funcionando aceptablemente sobre las muestras obtenidas de 5 personas. Se obtuvo un porcentaje de mediciones correctas del 78% para las palabras más extensas, bajando el rendimiento a un 70%, para las palabras pequeñas y dejando de funcionar correctamente para las

palabras con menos de 3 letras como por ejemplo las palabras "al, del, y", sobre todo para hallar la línea base superior debido a la poca cantidad de máximos.

Esta característica puede ser fácilmente contrastada por los peritos calígrafos, ya que la metodología que emplean para realizar esta medición también es medir la distancia ente las líneas bases.

Nuestra línea futura de investigación será incluir una mayor cantidad de escritores en la base de datos y trabajar para lograr la mayor robustez posible. Por un lado para abarcar gran variabilidad de estilos de escritura y por otro lado porque de la fiabilidad de la extracción de la característica dependerá la obtención de buenos resultados en la siguiente etapa de reconocimiento o clasificación de escritores, etapa que se abordará cuando se tenga una cantidad importante de características extraídas. Por este motivo también se deberá implementar la extracción de otras características que permitan identificar a las personas por medio de la escritura.

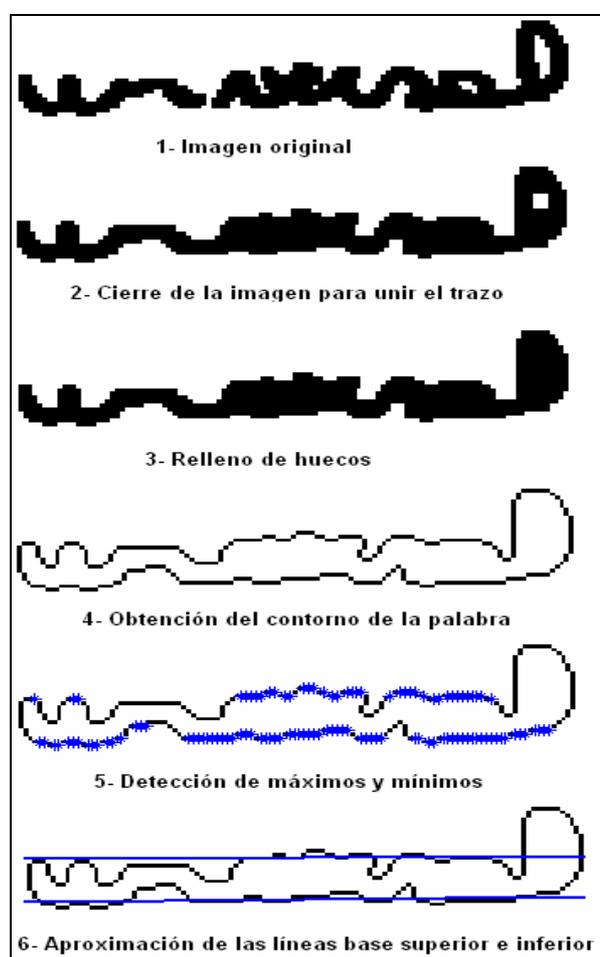


Fig. 6. Detección de las líneas base superior e inferior.

REFERENCIAS

- [1] Viñals Francisco y Puente M.^a Luz. "Pericia caligráfica judicial". Editorial Herder (2001).
- [2] Sargur N. Srihari, Sung-Hyuk Cha, Hina Arora y Sangjik Lee. "Individuality of Handwriting". J Forensic Sci, July 2002, Vol. 47, No. 4, Paper ID JFS2001227_474.
- [3] C. M. Grening, V. K. Sagar y C. G. Leedham. "Handwriting Identification Using Global and Local Features for Forensic Purposes". European Convention on Security and Detection, Mayo 1995, Conference publication N° 408, p. 272-278.
- [4] C. Tomai, B. Zhang y S. Srihari. "Discriminatory Power of Handwritten Words for Writer Recognition". Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04).
- [5] D. Valkaniotis, J. Sirigos, N. Antoniales y N. Fakotakis. "Text-Independent Off-Line Writer Recognition Using Neural Networks". ICECS '96, p. 692-695.
- [6] H. E. S. Said, G. S. Peake, T. N. Tan y K. D. Baker. "Writer Identification from Non-uniformly Skewed Handwriting Images". Proceeding of the 9th British Machine Vision Conference 1998, p. 478-487.
- [7] U.-V. Marti y H. Bunke. "Text Line Segmentation and Word Recognition in a System for General Writer Independent Handwriting Recognition". Sixth International Conference on Document Analysis and Recognition (ICDAR '01), september 2001, Seattle, Washington, p. 159-163.
- [8] H. E. S. Said, G. S. Peake, T. N. Tan y K. D. Baker. "Writer Identification from Non-uniformly Skewed Handwriting Images". Proceeding of the 9th British Machine Vision Conference 1998, bmva.ac.uk, p. 478-487.
- [9] N. Otsu. A threshold selection method from gray-level histograms. IEEE Transaction on Systems, Man and Cybernetics, 9(1):62-66, Jan. 1979.
- [10] Laurence Likforman-Sulem, Anahid Hanimyan and Claudie Faure. "A Hough Based Algorithm for Extracting Text Lines in Handwritten Documents". Internacional Conference in Document Analysis and Recognition, ICDAR'95 Montréal, p. 774-777.
- [11] I. Pitas. "Digital Image Processing Algorithms and Applications". John Wiley & Sons, Inc., 2000.
- [12] G. Nicchiotti y C. Scagliola. "Generalised Projections: a Tool for Cursive Handwriting Normalisation". International Conference on Document Analysis and Recognition, 8-1999, págs. 729-732.
- [13] Jaekyu Ha, R.M. Haralick, I.T. Phillips. "Document page decomposition by the bounding-box project," ICDAR, 1995, Vol. 02, N° 2, p. 1119.
- [14] Image Processing Toolbox for Matlab. User's Guide Version 5.
- [15] R.C. González y R.e. Wood, "Digital Image Processing". Addison-Wesley Publishing Company, Inc., 1993.
- [16] S. Madhvanath, G. Kim and V. Govindaraju. "Chaincode Contour Processing for Handwritten Word Recognition". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, N° 9, September 1999.
- [17] Carlos Manuel Travieso González. "Reconocimiento de Formas Manuscritas con Modelos Ocultos de Markov". Tesis Doctoral en el Grupo de Procesado de Señales Biológicas. U.L.P.G.C., mayo 2002
- [18] W. Chin, M. Harvey y A. Jennings. "Skew Detection in Handwritten Scripts". IEEE Region 10 Annual Conference. Speech and Image Technologies for Computing and Telecommunications (TENCON '97), Vol. 1, p. 319-322.
- [19] Adolfo Gustavo Suárez Lorenzo. "Segmentación de Texto Manuscrito". Proyecto de Fin de Carrera en el Grupo de Procesado de Señales Biológicas. U.L.P.G.C., diciembre 2001.