

Capítulo 7 Algoritmos de Segmentación

Capítulo 7 Algoritmos de Segmentación

7.1 Investigación y selección de algoritmos de segmentación de publicaciones recientes

El proceso de investigación de los algoritmos de segmentación involucró una búsqueda a través de Internet, sobre la cual se seleccionaron documentos provenientes de distintos centros de investigación de diversas partes del mundo. Existen algunos documentos que fueron descartados debido a que no fueron publicados a la comunidad científica a través de *papers* o que no presentaban una referencia sobre autores, origen o fecha en la que los proyectos fueron desarrollados.

A través de la lectura del material recopilado, nos dimos cuenta de que existen diversos tipos de segmentación de imágenes, basados en las características de las palabras. Existen escrituras que se caracterizan por usar letra tipo molde pero que entre los caracteres que la componen existe un traslape, es decir que uno se encuentra encima de otro. Por otra parte, se han desarrollado algoritmos que buscan separar letras impresas por una máquina pero que por diversos factores fueron encimadas. La Figura 7.1 muestra este tipo de diferencias.

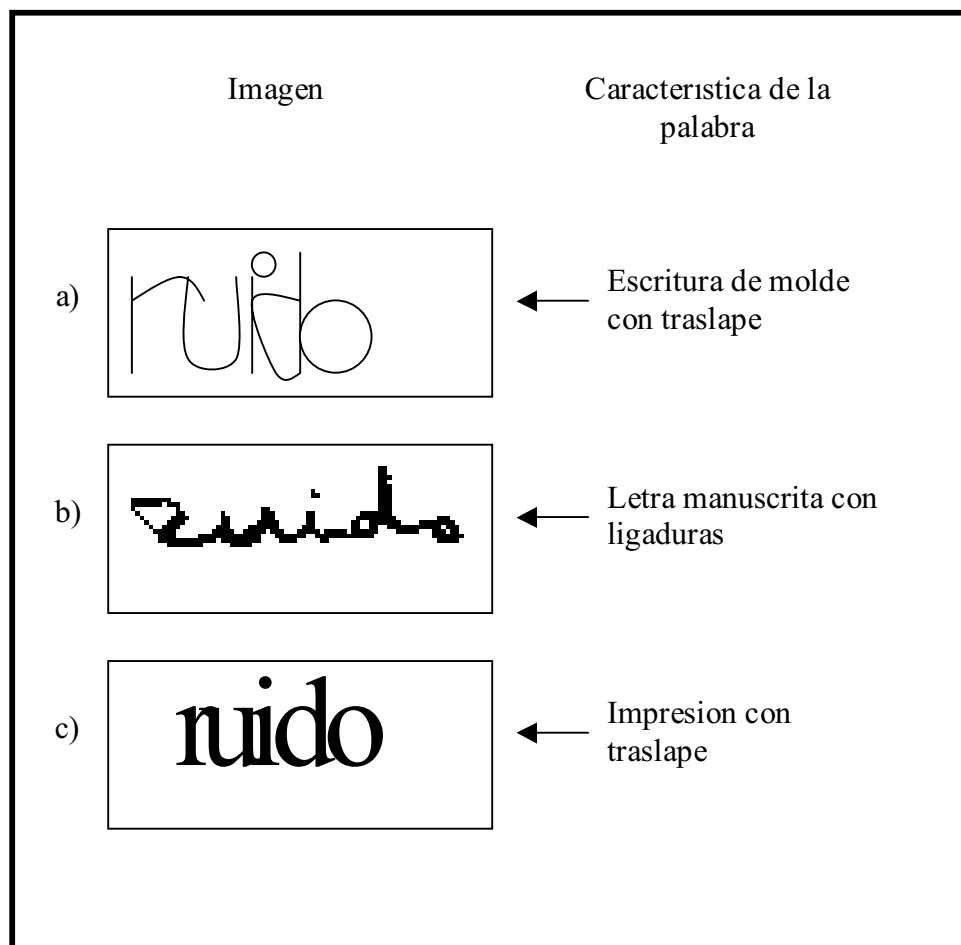


Figura 7.1 Tipos de imágenes sobre las cuales se especializan los algoritmos de segmentacion

Un impedimento que se encontró fue que algunas investigaciones no presentan sus resultados en el proceso de segmentación, sino que únicamente los muestran basándose en el reconocimiento como un sistema global. Esta desventaja no permitió comparar su eficiencia con respecto al resto de los métodos de segmentación. Un caso fue el encontrado en [Park & Govindaraju, 2000].

Cabe destacar que los porcentajes de éxito de cada algoritmo en muchas ocasiones depende del tipo de base de datos utilizada como entrada al proceso de segmentación. En la descripción del proyecto presentado por [Breuel, 1998], se utiliza una base de datos proveniente de formas de Censo de 1990 de los Estados Unidos. Este tipo de formas se caracteriza por poseer campos de escritura de un tamaño determinado, así como una línea horizontal sobre la cual se guía la persona. Debido a la importancia del documento, la tendencia de quienes lo llenan es el de escribir lo más legible posible recurriendo al tipo de letra descrita en el inciso a) de la Figura 7.1., la cual no resulta de interés para el proyecto. Además ciertos campos poseen pequeñas divisiones dentro de sí mismos para facilitar al usuario el escribir una letra dentro de cada casillero.

En el proyecto presentado por [Kavallieratos & Antoniadis, 1997] la base de datos proviene de formas de aplicación de una compañía de seguros. El tipo de letra plasmada es más similar a la que será utilizada en este proyecto (cuyas características se detallan en el Capítulo 3), sin embargo las imágenes originales presentan un porcentaje de ruido y daños mucho menores.

Debido a la gran variabilidad de resultados que presentan los algoritmos de segmentación a partir de la base de datos de entrada, se decidió seleccionar aquellos cuyos resultados se dieron a partir de la base de datos CEDDAR [Debashish & Sargur, 1995]. Esta se caracteriza por tener una gran variedad de tipos de letra manuscrita provenientes de sobres de la Oficina de Correo Postal de los Estados Unidos. De esta manera se pudo

determinar el alcance de cada algoritmo. En el Apéndice C se pueden observar los porcentajes de éxito de algunos de éstos.

7.1.1 Algoritmos de segmentación desechados

Uno de los algoritmos descartados para este proyecto es el presentado por [Breuel, 1998]. Su esencia es el establecimiento de líneas de segmentación a lo largo de los contornos verticales de cada letra en la palabra. La sobresegmentación es permitida a este nivel y se presenta como solución el uso de un grafo de hipótesis sobre el cual se decidirá la combinación correcta de segmentos.

A pesar del éxito que describe para letras de molde traslapadas, su éxito en la escritura manuscrita con ligaduras es muy bajo (según se explica dentro del mismo), por tanto no resultó conveniente para este proyecto. Otra razón fue el no presentar resultados en base a la segmentación.

Algunos otros algoritmos no fueron utilizados debido a que requerían un conocimiento especializado sobre un área de estudio. Este fue el caso del método descrito en [Maragoudakis & Fakotakis, 2001]. Los resultados que presentan son bastante satisfactorios sin embargo, el análisis e implementación requerirían bastante tiempo ya que se necesitaría tener conocimientos sobre máquinas vectoriales de apoyo y sobre elucidación probabilística (conceptos básicos de este método). Por tanto su desarrollo no resultó factible para el plazo establecido para el proyecto. Se optó por algoritmos con resultados similares que requerían menor tiempo de estudio.

En las secciones siguientes se hablará sobre aquellos algoritmos que si fueron desarrollados y las razones por las cuales fueron escogidos.

7.2 Algoritmos Utilizados en la aplicación HAWOST

Dos algoritmos heurísticos de segmentación fueron desarrollados, el método de densidad de píxeles con calculo de ancho de pincel y el algoritmo de huecos blancos. Un tercer método fue desarrollado que consiste en utilizar alguno de los métodos heurísticos anteriores ayudado por el asesoramiento de una RNA (algoritmo híbrido). En las siguientes se secciones se explicarán cada uno de éstos.

7.2.1 Algoritmo de detección de huecos blancos

El objetivo de este método es encontrar huecos blancos dentro de la palabra. Estos huecos corresponden a letras como la “a”, “e”, “o”, “p”, “q”, “g”, “d”, “g” o “b”. La detección de estas zonas nos permite saber que existe un punto de segmentación antes y después del hueco [Blumenstein & Verma, 1999].

Primero se recibe una matriz binaria (unos y ceros) correspondiente a la imagen analizada. A continuación se generará una segunda matriz que contendrá aquellos píxeles blancos que se encuentren rodeados por píxeles negros. Para lograr lo anterior se hace un recorrido de cada celda con valor 0 de la matriz inicial, verificando si existe al menos un

píxel negro arriba y debajo dentro de la columna donde se encuentra localizada, y si existen otros a su izquierda y derecha (dentro su renglón). Si cumple todos estos requisitos se coloca el valor de uno en la posición i, j en la matriz secundaria. Una vez obtenida esta segunda matriz el siguiente paso es generar un vector que indique a partir de qué posición comienza el hueco y cuando termina. Por último se establecerán como puntos de segmentación aquellas coordenadas en el eje horizontal donde inicia y termina un hueco añadiéndole un número fijo de traslación que permitirá cubrir la zona de píxeles negros que rodean al hueco, ver figura 7.2.

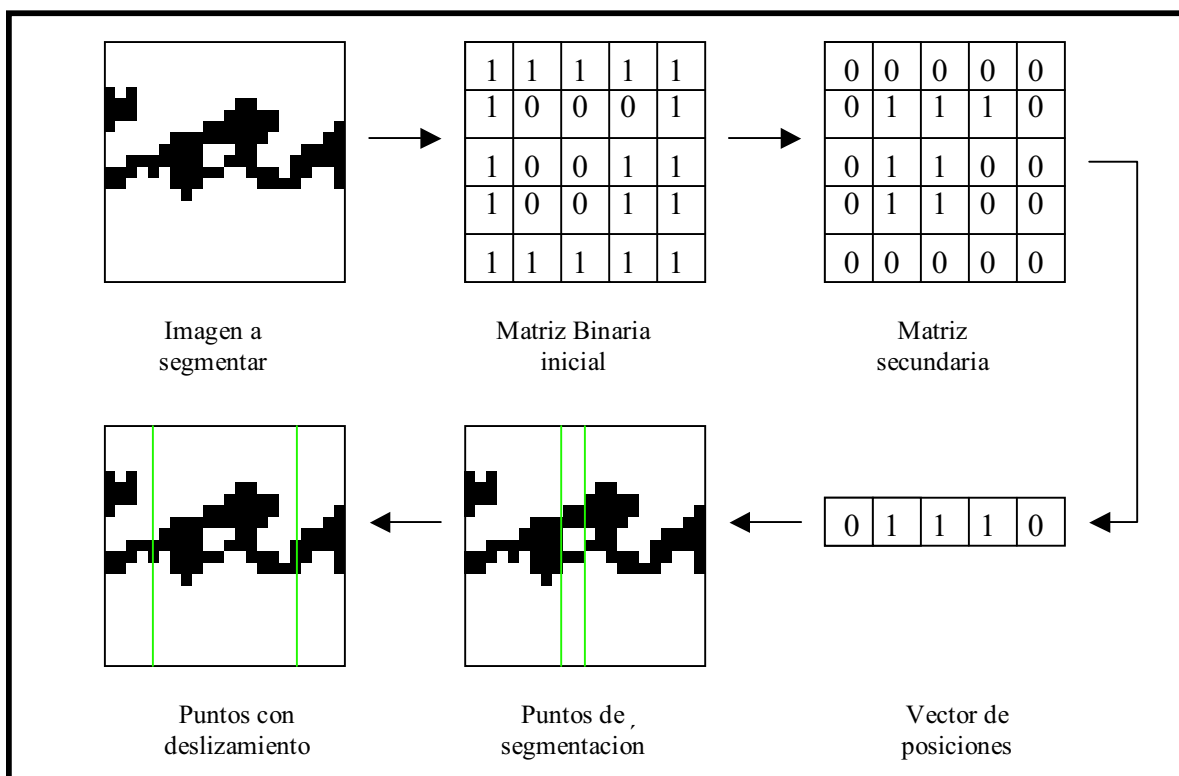


Figura 7.2 Método de segmentación por huecos blancos

7.2.2 Algoritmo de densidad de píxeles con cálculo de ancho de pincel

El procedimiento comienza con un conteo del número de unos en cada columna de la matriz binaria que representa la imagen de la palabra. En base a este conteo se elabora un histograma donde cada barra representa estas sumas. A continuación se elabora un segundo histograma denominado histograma con umbral. Para construirlo se define un número, si la barra del histograma inicial lo rebasa, entonces adquiere un valor de 10 en el segundo histograma y si es inferior se le asigna el valor de cero. Esto nos genera un histograma donde se observan bloques, cada uno representa una letra y cada “valle” entre ellos representa una ligadura. Los puntos de segmentación se colocarán a la mitad de cada valle, ver figura 7.3.

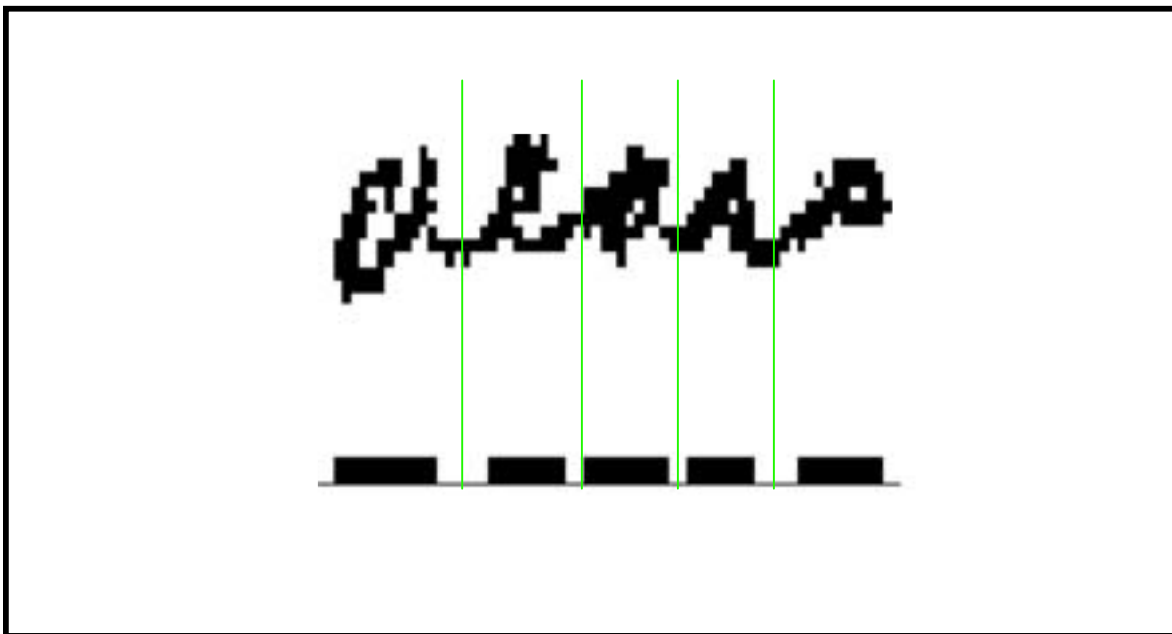


Figura 7.3 Algoritmo heurístico de densidad vertical de píxeles

Como se había mencionado en capítulos anteriores, las palabras del Gral. Porfirio se caracterizan por presentar variaciones de intensidad y ancho de tinta entre una y otra. Así que para mejorar este algoritmo se desarrolló un proceso de detección de ancho de pincel para que en base a este valor, se estableciera el umbral que determina si la densidad de un histograma corresponde a una letra o a una ligadura.

Para facilitar la comprensión de esta necesidad se presentan dos tipos de imágenes en la Figura 7.4. Como se observa, en la primera existe un ancho de tinta bastante grueso. Así que el umbral a establecer para diferenciar un histograma correspondiente a una letra o a una ligadura deberá ser alto. Por el contrario en la imagen de la derecha se observa una palabra con ancho de pincel muy delgado y que necesita de un umbral bajo.

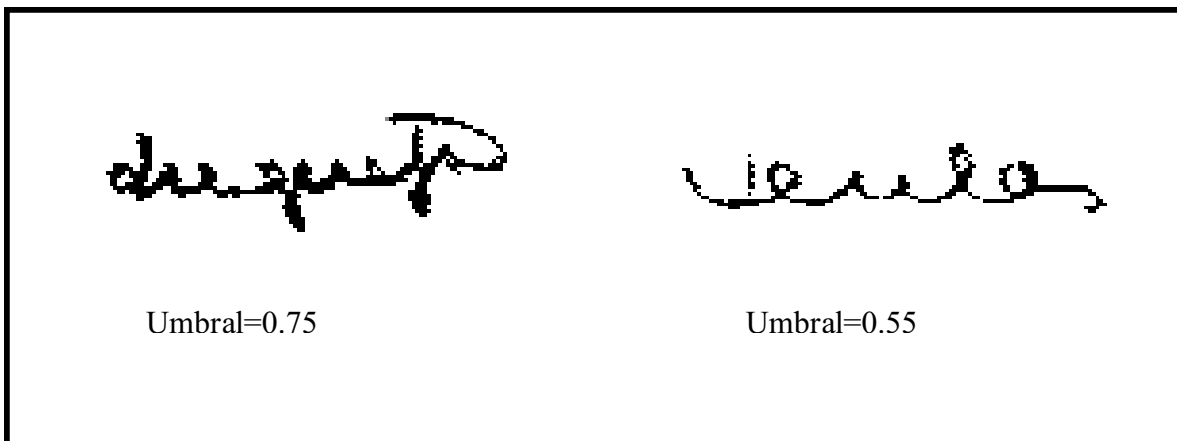


Figura 7.4 Umbrales apropiados para palabras con ancho de tinta variable

7.2.3 Algoritmo de segmentación Híbrido

El algoritmo presentado en [Blumenstein & Verma, 1999] fue uno de los más apropiados para la segmentación de la escritura del Gral. Porfirio Díaz. La ventaja que presenta es el englobar varios métodos de segmentación heurísticos con una red neuronal artificial asesora. Dicha ventaja radica en que según la experiencia obtenida del proyecto anterior, un solo método de segmentación no es suficiente para particionar las palabras de nuestra base de datos y por otra parte es necesario establecer un mecanismo de eliminación de la sobresegmentación para alcanzar niveles óptimos. Como se puede observar en el Apéndice C, los porcentajes de éxito de este algoritmo sobre la base de datos CEDDAR son bastante altos.

Los algoritmos heurísticos que utiliza son: la detección de hoyos blancos y la proyección de la densidad vertical de píxeles. Como primer paso, es necesario establecer el promedio de ancho de letra de las palabras de la base de datos. La herramienta HAWOST proporciona una manera fácil de determinar este valor. El procedimiento es necesario realizarse una sola vez. El usuario debe de segmentar de manera manual la imagen desplegada en la aplicación. Una vez hecho esto, automáticamente se determinará el promedio de ancho de letra que posteriormente servirá como apoyo a los algoritmos heurísticos para eliminar aquellos segmentos catalogados como letra pero que no se acercan al promedio de ancho.

El método híbrido consiste en aplicar el algoritmo de huecos blancos y de densidad de píxeles para sobresegmentar la imagen, de esta manera nos podemos asegurar que los puntos de segmentación correctos ya se encuentran establecidos y lo único necesario es eliminar los incorrectos. Para eliminar los puntos mal establecidos se aplica una RNA previamente entrenada (ver sección 6.3) que se encargará de catalogarlos y eliminarlos. Usando esta técnica se obtuvieron palabras mejor segmentadas como lo muestra la figura 7.5.

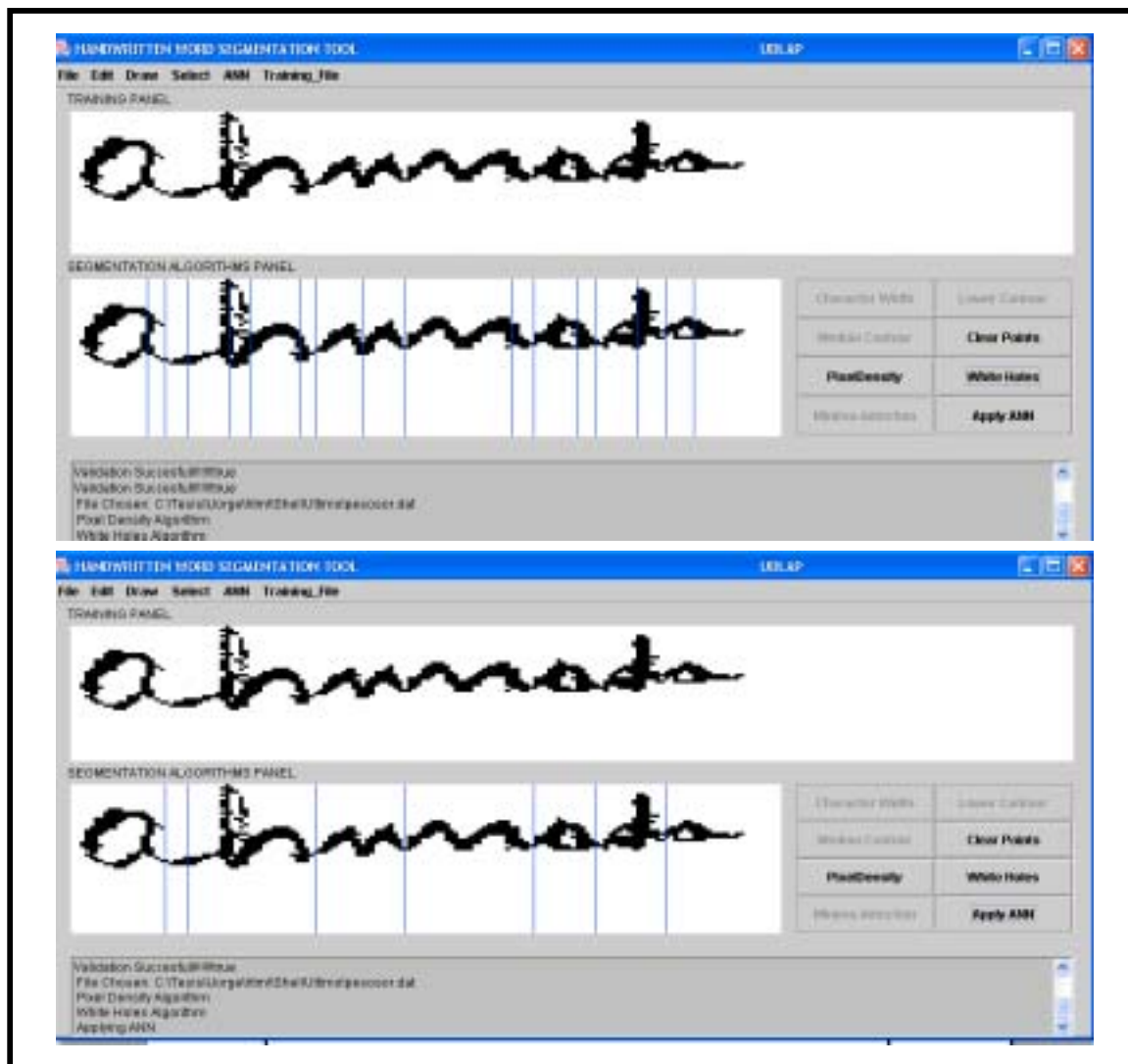


Figura 7.5 Método de segmentación híbrido

En la siguiente sección se mostrarán resultados cuantificables sobre los resultados obtenidos al aplicar estos métodos sobre las imágenes extraídas de los manuscritos del Gral. Porfirio Díaz