

Clasificación de Números Manuscritos Basados en Prototipos Creados Usando La Distancia Euclidiana como Umbral

G. CEBALLOS B.
gceballos@ing.uchile.cl
Departamento de Ingeniería Eléctrica
Universidad de Chile
Tupper 2007, Casilla 412-3, Santiago.
CHILE

1. INTRODUCCION

En este proyecto se aplicó un método de clasificación de números manuscritos basado en prototipos creados usando la distancia euclidiana o error cuadrático medio como umbral de clasificación. La base de datos utilizada consta de 500 imágenes de números (de 20x20 pixeles) para entrenamiento y 200 para validación con sus respectivas salidas etiquetadas entre 0 y 9 obtenidas de la base de datos MNIST. También, este método fue comparado con un algoritmo de redes neuronales perceptrón multicapa (MLP) con entrenamiento supervisado backpropagation y funciones de activación sigmooidal tipo “logsig”. La mejor clasificación lograda con esta base de datos fue de 80%. Este porcentaje de clasificación correcta fue bastante similar al obtenido con la red neuronal (82% aproximadamente en el mejor caso).

2. METODO

El método utilizado para el reconocimiento de números manuscritos consta básicamente de tres estados. En la Figura 1 se muestra dicho esquema.

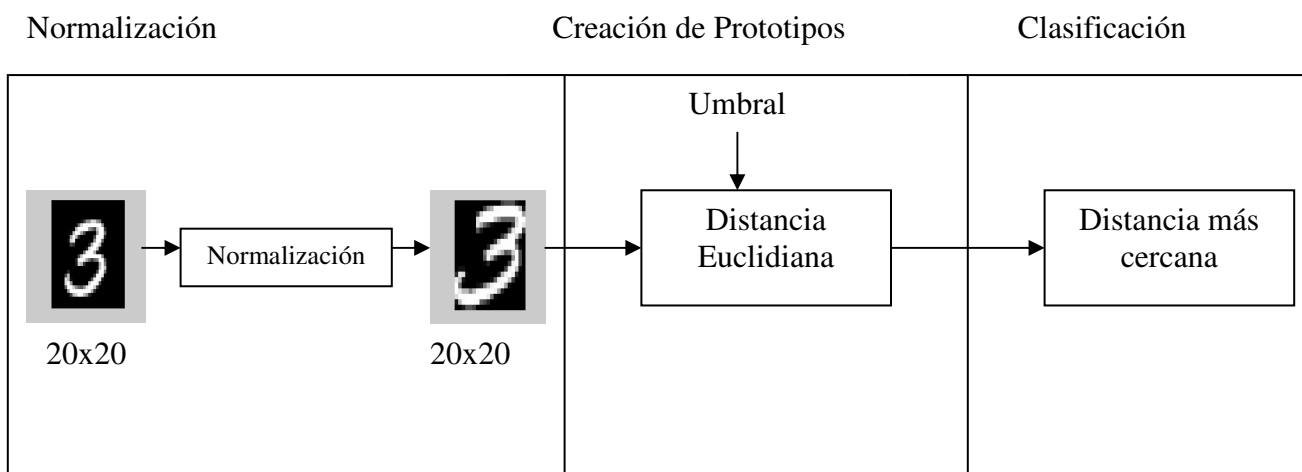


Figura 1: Sistema de Reconocimiento de Dígitos Manuscritos basados en prototipos

En el primer bloque, el número es expandido de tal manera que el dígito ocupe toda la imagen, teniendo cuidado de preservar su razón de aspecto, es decir, la razón entre la altura y ancho del dígito debe permanecer relativamente inalterable (normalización).

Luego, en el segundo bloque, los dígitos son prototipados en base a la distancia euclidiana que exista entre los pixeles del número patrón que se presenta al sistema y los del prototipo más cercano existente (si es que este fuera el caso). La primera vez que inicia el algoritmo, el primer dígito o patrón que se le presenta al sistema es a su vez el primer prototipo creado. Posteriormente, al segundo dígito se le calcula la distancia euclidiana con respecto a los prototipos ya creados (hasta el momento solo hay 1 prototipo creado). Entonces, si esta distancia es menor que cierto umbral o distancia de aceptación, se chequea que el dígito pertenezca a la clase del prototipo y en caso afirmativo, se promedia con dicho prototipo. En el caso contrario (que la distancia entre el dígito presentado y el prototipo sobrepase dicho umbral pero que el dígito sea de la misma clase que el prototipo, se crea un nuevo prototipo para dicha clase siendo éste, el dígito presentado. Ahora bien, si la distancia es menor que dicho umbral pero el dígito es de diferente clase que el prototipo, se crea también, un nuevo prototipo para la clase del dígito. Es aquí donde se le introdujo una variante al método en el sentido que para la etapa de creación de prototipos, no es necesario realizar la comparación de distancia con todos los prototipos existente para un dígito patrón, sino que basta con comparar dicha distancia con los prototipos de su clase solamente con lo cual el método mejora bastante la rapidez de creación de prototipos. No obstante lo anterior, para la etapa de clasificación si será imperioso que el dígito se compare, en el sentido de la distancia euclidiana, con todos los prototipos previamente creados.

A continuación, se muestra la ecuación para el cálculo de la distancia euclidiana .

$$d(w_m^{ck}, z) = \sqrt{\sum_{i=1}^{20} \sum_{j=1}^{20} (w_m[i][j] - z[i][j])^2} \quad (1)$$

En que “z” es la imagen del dígito de 20x20 pixeles y Wm es el prototipo de la clase ck con el cual se esta calculando dicha distancia.

Por otro lado, la ecuación para la modificación (promediación) de un prototipo en caso que dicho dígito posea una distancia menor al umbral y sea de la misma clase que su prototipo más cercano es la siguiente:

$$w_m[i][j] = \frac{w_m[i][j]Nprot(m) + z[i][j]}{Nprot(m) + 1} \quad (2)$$

donde $N_{\text{prot}}(m)$ corresponde al número de patrones presentados previamente que han contribuido al prototipo w_m de la clase respectiva.

A modo de ejemplo, se muestra a continuación la creación de 75 prototipos para un umbral de distancia de 8 (en este caso los dígitos están representados en niveles de grises normalizados entre 0 y 1).

$$N^{\circ} \text{Prototipos de la Clase} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 9 \\ 1 \\ 14 \\ 3 \\ 11 \\ 10 \\ 5 \\ 5 \\ 11 \\ 6 \end{bmatrix}$$

Hay algo interesante que notar, en el caso de un umbral=8, para la clase 1 (equivalente al dígito manuscrito “1”), todos los dígitos patrones de dicha clase contribuyeron a crear un único prototipo que representa dicha clase. Dicho de otra manera, para este umbral, todos los dígitos N°1 son muy similares. Esto es al revés de lo que sucede con el dígito N°3 (clase 3) en que se crearon 14 prototipos.

También es necesario notar que mientras más pequeño sea el umbral, mayor será la cantidad de prototipos creados para cada clase. Esto se ve a continuación, en que en este caso la distancia umbral es de 5.

$$N^{\circ} \text{ Prototipos de la Clase} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 35 \\ 5 \\ 46 \\ 50 \\ 51 \\ 49 \\ 56 \\ 38 \\ 50 \\ 41 \end{bmatrix}$$

En este caso, se crearon 421 prototipos.

Para el caso límite en que el umbral sea cero, se crean tantos prototipos como dígitos existan para al etapa de entrenamiento o mejor dicho, etapa de creación de prototipos.

3. RESULTADOS

Tabla N°1: Desempeño en la clasificación de prototipos para diversos umbrales

Umbral de distancia para la creación de Prototipos	N° de Prototipos creados	% de Clasificación Correcta
8	75	76.0
7	191	79.5
6	329	77.5
5	421	80.0
4	460	80.0
3	477	80.0
2	493	80.0
1	500	80.0

Ahora bien, comparemos estos resultados con los obtenidos para una Red Neuronal Multicapa tipo Perceptrón (MLP).

Tabla N°2: Porcentaje de Clasificación Correcta de Dígitos Manuscritos usando una Red Multicapa tipo Perceptrón.

N° de simulación	Topología	Entrenamiento [%]	Validación [%]
1	(400):80:80:10	100	82.5
2		100	81.5
1	(400):50:50:10	100	80.5
2		100	74.5
1	(400):20:20:10	98.4	72
2		100	78.5
1	(400):80:10	100	75
2		100	79.5

Es evidente entonces, que para un umbral de 5, el método de clasificación de números manuscrito es comparable al desempeño de una red neuronal, siendo este último, bastante más lento en la etapa de entrenamiento.

A continuación, se muestra la matriz de confusión obtenida para el caso del umbral 8 en que se crearon 75 prototipos en total.

Tabla N°3: Matriz de confusión para un umbral de 8.

Dígito	0	1	2	3	4	5	6	7	8	9
0	88	0	4	4	0	0	4	0	0	0
1	0	77.8	0	0	11.1	0	5.55	5.55	0	0
2	0	0	85	0	5	0	0	0	10	0
3	0	0	8.7	82.6	0	8.7	0	4.35	4.35	0
4	0	5.88	0	0	64.71	0	0	17.65	0	11.76
5	0	0	0	0	6.25	50	6.25	0	37.5	0
6	0	0	0	0	0	0	100	0	0	0
7	0	0	0	0	15.79	0	0	73.68	0	10.53
8	0	0	5	5	5	5	0	0	80	0
9	0	0	8	4	4	8	0	8	4	64

De esta matriz se puede obtener información valiosa. Por ejemplo, es evidente que el dígito 5 fue el que más tuvo confusión, en particular, se confundió muchas veces con alguno de los prototipos del número 8. A su vez, otros dígitos que tuvieron alta confusión

fueron el 4 y el 7 los cuales se confundieron entre sí (aproximadamente en un 16% de los patrones).

La importancia de esta matriz radica en el hecho que a través de su análisis uno puede inferir donde es necesario acentuar más la búsqueda de nuevos tipos de prototipos o características que estos debiesen tener. Por ejemplo, como el dígito 5 se confunde mucho con el prototipo 8 uno debiese optar por mirar características geométricas de los dígitos, de tal manera de fijarse que el dígito 8 es cerrado arriba y abajo, por lo tanto aunque un 5 este muy cercano a un prototipo 8 en el sentido de la distancia euclidiana, este no debiese clasificarse en esa clase ya que no tiene cierre abajo y menos arriba.

Por último vemos que el dígito que menos se confundió fue el dígito 6, es decir, este dígito es muy separable al menos en el sentido de la distancia euclidiana, con los otros dígitos. Dicho de otro modo, este dígito resultó ser muy distinguible de los demás.

4. EJEMPLOS DE CONFUSION OBTENIDOS EN LA CLASIFICACION PARA UMBRAL 8



Figura 2: Ejemplo de confusión: (a) Dígito presentado en la clasificación y (b) Prototipo seleccionado.

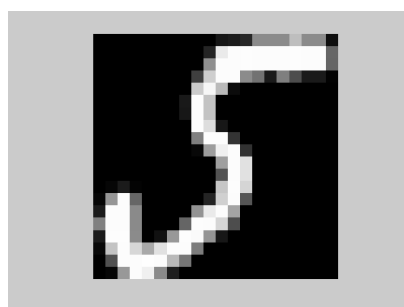


Figura 3: Un prototipo de la clase 5 no seleccionado para el dígito presentado en la Figura 2.

A continuación presentamos otro ejemplo de confusión que se presentó en la clasificación.



Figura 4: Otro ejemplo de confusión: (a) Dígito presentado en la clasificación y (b) Prototipo seleccionado.

En el caso de Dígito N°3 se crearon sólo 3 prototipos, esto implica que los números de la base de datos para entrenamiento (creación de prototipos en nuestro caso), eran bastante similares entre sí, no obstante, en la base de datos para validación, algunos de estos dígitos eran más parecidos a los 8 que a los 3 (en el sentido de la distancia euclidiana por supuesto) y es lo que se aprecia en la Figura 4. A continuación se muestran los tres prototipos creados para el dígito N°3 con umbral 8.

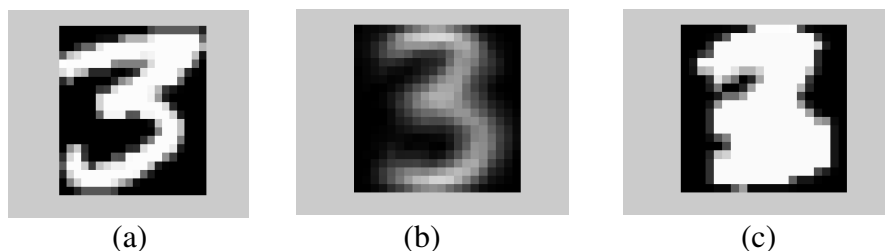


Figura 5: (a), (b) y (c) Prototipos creados para el N°3.

En el caso anterior, se ve que el tercer prototipo del dígito 3 es el más promediado por eso aparece mas denso, es decir, muchos dígitos 3 contribuyeron a crear ese prototipo.

Finalmente se muestra un último ejemplo obtenido con una porción mayor de la base de datos MNIST en donde los dígitos son de 28x28 pixeles y no están normalizados. En este caso, se usaron 2.000 dígitos para la creación de prototipos y otros 1.000 dígitos distintos para las pruebas de clasificación. Los resultados obtenidos fueron los siguientes.

Umbral de distancia para la creación de Prototipos	N° de Prototipos creados	% de Clasificación Correcta
8	129	90.4
7	346	92.5
6	795	93.8
5	1.341	94.2

Estos resultados son bastante comparables si se los compara con redes perceptrón multicapa [1]. Además, existen algunas mejoras que se pueden introducir para aumentar en 2 puntos porcentuales aproximadamente los porcentajes de clasificación. A saber, estas son (a) Votación entre los prototipos cuya distancia euclidiana al dígito para clasificación sea menor a una cierta distancia de aceptación, es decir, se elige la clase del prototipo que más veces apareció y (b) en caso que exista empate en la votación, se puede recurrir a una nueva mejora la cual incorpora reglas difusas (basada en la geometría de los dígitos anteriormente mencionada).

5. CONCLUSIONES

Se ha probado un método para clasificar números manuscritos basado en la creación de prototipos usando el concepto de distancia euclidiana o error cuadrático medio. Los resultados obtenidos son comparables a los que se obtienen con redes neuronales perceptrón multicapa (MLP). Además, la complejidad es mucho menor que la usada en la clasificación con redes neuronales. El tiempo de clasificación para el caso de la base de datos con 500 patrones para creación de prototipos y 200 patrones para clasificación (75 prototipos) fue de 8 segundos con un programa computacional implementado en MATLAB en un Pentium III de 350 Mhz y 128 Mbytes en RAM. En el caso en que el umbral se redujo a 5 (421 prototipos) este tiempo de clasificación fue del orden de los 47 segundos.

6. BIBLIOGRAFIA

[1] Lee, D.S. Srihari, S.N. and Pawlicki, T., "Experiments with Neural Network Models for Handwritten Digit Recognition", en "Systems and Signal Processing", R.N. Madan, N. Viswanadham, R.L. Kashyap (Eds.). pp. 757-774, 1991.

[2] Haykin, S., "Multilayer Perceptrons", en "Neural Networks: A Comprehensive Foundations", John Griffin (Ed), IEEE Press, Macmilan College Pu Co., pp 138-229, 1994.

[3] Perez CA, Held CA, Mollinger P, "Improvements on Handwritten Digit Recognition by Genetic Selection of Neural Network Topology and by Augmented Training", Proc. 1997 IEEE International Conference on Systems, Man and Cybernetics, Orlando USA, pp. 1487-1491, Oct.12-15,1997.