

Tarea 1 – Reconocimiento de Patrones

Rodrigo Nazar Meier, *Pontificia Universidad Católica de Chile*

Abstract—En el reconocimiento de patrones es popular el problema de reconocimiento de caracteres en variadas fuentes. Su popularidad se debe a la infinidad de aplicaciones que tiene este tipo de trabajos en la realidad: poder detectar letras y números escritos a mano (transcripción de textos), reconocer información de objetos en la vida real que portan caracteres alfanuméricos, entre muchas otras. La presente tarea muestra una forma simple de clasificar las letras A, S, D, F y G en distintos tipos de fuente y con rotaciones.

Index Terms—Reconocimiento de Patrones, Reconocimiento de Caracteres, Momentos de Hu.

1 SOLUCION PROPUESTA

Las imágenes disponibles para el *training* correspondían a representaciones monocromáticas en distintas fuentes de las letras A, S, D, F y G. Algunas letras estaban rotadas, otras de tamaños variables y también no necesariamente estaban centradas, es por esto que para poder clasificarlas necesitamos obtener características que sean **invariantes a la rotación, escala y traslación**.

Las características que elegí obtener para lograr a cabo esta tarea son los Momentos de Hu [1]. Este tipo de características poseen todas las invarianzas listadas anteriormente. Cabe recalcar que en este experimento sólo necesitamos enfocarnos en las características geométricas de las letras que buscamos clasificar, ya que estas se diferencian por este medio y no por cambios de intensidad o de texturas.

2 EXPERIMENTOS REALIZADOS

Para los experimentos se partitionaron las imágenes entregadas en dos distintas categorías: *training* y *testing*, el primero contando con un *dataset* de 72 imágenes por clase y el segundo de 36 por clase. Con estos datos, se procede a obtener los Momentos de Hu de cada una de las imágenes. Estos se y se almacenan en dos archivos del tipo *json*, cada uno para cada categoría. Luego se normalizan los valores de cada momento de hu mediante una normalización *Min Max*.

Una vez ya generados los archivos de datos, se procede a clasificar cada letra del *testing*. Esto se llevó a cabo calculando la distancia euclidiana del vector de momentos de la imagen de *testing* con todos los datos de *training*, luego se tomó una muestra de los 5 elementos con menor distancia y finalmente se seleccionó la clase que más se repetía en la muestra.

Para mejorar el desempeño del clasificador, se hizo un ciclo iterativo en donde se calcula el porcentaje de aciertos del clasificador por cada combinación distinta de los Momentos de Hu. Esto permitió deshacerse de información redundante o que no aportaba en el clasificador.

Los mejores resultados (utilizando la menor cantidad de momentos) se obtuvieron utilizando sólo los primeros tres Momentos de hu, dando así un 100% de aciertos con todas las imágenes de *testing*. Los resultados se muestran en la Fig. 1.

	Valores predichos				
	A	S	D	F	G
A	36	0	0	0	0
S	0	36	0	0	0
D	0	0	36	0	0
F	0	0	0	36	0
G	0	0	0	0	36
Total	100%	100%	100%	100%	100%

Fig. 1: Matriz de confusión

Por su parte, en la Fig. 2 se aprecian los histogramas de los momentos utilizados. A simple vista no es fácil observar como el clasificador separa las clases completamente en dos dimensiones, pero en la representación tridimensional del diagrama de dispersión se logra ver una tendencia de separación.

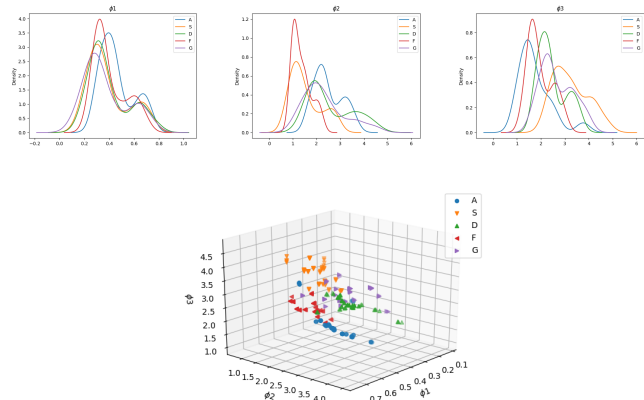


Fig. 2: Diagramas de dispersión de los momentos de hu utilizados.

Importante: revisar el archivo *README.md* de la tarea para ver detalles de la implementación.

3 CONCLUSIONES

Los Momentos de Hu son robustos y simples descriptores geométricos, que le otorgaron un perfecto desempeño al clasificador. Para futuras implementaciones sería interesante conocer su desempeño con un conjunto de clases más numeroso.

REFERENCES

- [1] Ming-Kuei Hu, "Visual pattern recognition by moment invariants," in *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179-187, Febrero 1962.