

CAPÍTULO 2

MARCO TEÓRICO

Reconocimiento de Patrones

El reconocimiento de patrones es utilizada actualmente para la solución de tareas tales como el reconocimiento de caracteres, de huellas digitales y reconocimiento del habla, entre muchos otros. Por esta razón, desde hace varios años se han venido desarrollando diferentes metodologías que intentan resolver este tipo de problemas.

Metodologías para el reconocimiento de patrones

De manera general, podemos decir que existen tres metodologías básicas para el reconocimiento de patrones [2]: las heurísticas, las matemáticas y las lingüísticas.

Heurísticas

Esta metodología es la que hace uso de la experiencia y la intuición humana. Por lo general, los sistemas que son desarrollados bajo estos métodos, están hechos a la medida del problema que se desea resolver.

Matemáticas

Este tipo de metodologías hacen uso de las propiedades comunes de los patrones y se basan en las reglas de clasificación formuladas en un marco matemático. Además, este enfoque se divide a su vez en otras dos categorías: las determinísticas y las estadísticas. Entre los métodos determinísticos podemos encontrar, por ejemplo, los algoritmos de aprendizaje iterativo tales como los algoritmos utilizados para el entrenamiento de redes neuronales. Por otra parte, dentro de los métodos estadísticos podemos encontrar a los clasificadores basados en la regla de clasificación de Bayes.

Lingüísticas (sintácticas)

Por último encontramos las metodologías lingüísticas en las cuales se hace uso de los elementos primitivos que componen a los patrones (sub-patrones) y la relación que existe entre ellos.

Etapas para el reconocimiento de patrones

De manera general, cuando se intenta resolver algún problema que implique el reconocimiento de patrones se deben de tomar en cuenta las siguientes etapas:

- Adquisición de datos.
- Preprocesamiento.
- Extracción de características.
- Reconocimiento.

Adquisición de datos

Para poder realizar el reconocimiento es necesario, realizar o implementar alguna etapa de adquisición de los datos que describen al patrón que se desea clasificar. Esto implica que, el método que se utilizará para realizar la adquisición de los datos, dependerá de la naturaleza del patrón. Por ejemplo, si se desea desarrollar un sistema para reconocer caracteres, esta etapa se podría realizar con la ayuda de una cámara digital o utilizando un escáner.

Preprocesamiento

Una vez que ya se cuenta con los datos que describen a cada uno de los patrones, es más conveniente realizar una etapa de preprocesamiento sobre cada uno de ellos en lugar de ser dados como entrada del sistema tal y como fueron obtenidos durante la etapa de adquisición de datos.

El realizar un preprocesamiento sobre los datos tiene varias ventajas; la principal de ellas es que puede reducir la dimensionalidad de los datos, lo cual mejora substancialmente la

ejecución del sistema, sobre todo cuando se utiliza una metodología como la de redes neuronales.

Dentro de esta etapa se puede preprocesar el patrón de entrada de tal forma que todos los patrones tengan el mismo tamaño (escala) consiguiendo con esto que el sistema sea invariante al escalamiento. Además de esto, también se busca lograr que el sistema sea invariante a la traslación. Cuando un sistema es invariante a la traslación y al escalamiento de los patrones, se dice que el sistema cuenta con un conocimiento previo.

Extracción de características

Uno de los principales problemas en el reconocimiento de patrones, es encontrar una manera óptima de representar la información original que describe a cada uno de los patrones. Este problema es conocido como extracción de características. Este proceso de extracción de características trata de reducir la cantidad de información (reducción de dimensionalidad) que representa a cada uno de los patrones, obteniendo de esta forma, un vector de características que represente de la mejor manera posible al patrón original.

La extracción de características debe de cumplir con las siguientes condiciones:

- La dimensionalidad del vector de características debe de ser menor que la del patrón original.
- Las características deben representar una codificación óptima de la entrada, perdiendo la información que no sea muy importante [3].

Especialmente, cuando se realiza el reconocimiento de patrones utilizando redes neuronales, la extracción de características debe tratar de obtener un vector de características con una dimensionalidad mucho menor a la del patrón original, puesto que, un vector con una dimensionalidad más pequeña que sea dado como entrada a la red neuronal tiene varios beneficios. En primer lugar, la cantidad de pesos que deben de ser aprendidos es menor, y en segundo lugar, al tener menos pesos, el tiempo de entrenamiento puede ser reducido considerablemente.

Reconocimiento

Finalmente, una vez que se han realizado las etapas de preprocesamiento y extracción de características, se procede a realizar el objetivo principal del sistema: el reconocimiento. De manera general, la etapa de reconocimiento consiste en obtener una salida del sistema como respuesta a un conjunto de estímulos dados como entrada (patrones).

Redes Neuronales

Las redes neuronales o modelos conexionistas como también se les conoce, están inspiradas en el comportamiento del cerebro humano. Esto significa, que cuando se desarrolla un modelo de red neuronal se trata de imitar el comportamiento del cerebro humano. Una red neuronal está compuesta de un conjunto de elementos que imitan a los componentes neuronales biológicos, tan es así, que dichos elementos son unidades de procesamiento llamados neurones artificiales, los cuales tienen conexiones de entrada (dendritas) que provienen de otros neurones y conexiones de salida (axón) hacia otros neurones. Para poder obtener un buen desempeño, los neurones de una red neuronal están masivamente interconectados con otros neurones mediante la serie de conexiones tanto de entrada como de salida, las cuales tienen un cierto peso.

De manera general las redes neuronales presentan tres principales características [4]: aprendizaje, generalización, adaptabilidad. La característica de aprendizaje se refiere a que al igual que el cerebro humano, una red neuronal tiene la capacidad de almacenar conocimiento mediante un proceso de aprendizaje (entrenamiento). Este conocimiento es almacenado por los pesos de las conexiones entre los neurones que conforman a la red neuronal. Por otra parte, la capacidad de generalización significa que se pueden obtener salidas razonables cuando se usan entradas diferentes a las utilizadas durante el proceso de entrenamiento. Por último, la característica de adaptabilidad significa que una red neuronal puede ser reentrenada para funcionar adecuadamente ante cambios en su medio ambiente.

Existen diferentes arquitecturas de redes neuronales, entre las cuales podemos mencionar: redes alimentadas hacia delante de capa simple, redes alimentadas hacia delante multicapas,

y redes recurrentes. Una de las arquitecturas más utilizadas para el reconocimiento de rostros es la red alimentada hacia adelante multicapas o red de perceptrones multicapas (MLP).

Básicamente, este tipo de arquitectura de red consta de tres componentes principales: una capa de entrada, una o más capas ocultas, y una capa de salida. Este tipo de redes comúnmente son entrenadas utilizando el algoritmo de retro-propagación de error el cual consta de dos fases. La primera, es una fase hacia adelante, en la cual la información colocada en los neurones de la capa de entrada (patrón) es propagada hacia adelante a través de las capas ocultas y hasta la capa de salida, la cual genera en los neurones que la conforman, la respuesta a la información dada como entrada. Durante esta fase, todos los pesos de las conexiones entre neurones son utilizados. La segunda fase es conocida como fase hacia atrás, la cual consiste en modificar los valores de los pesos de acuerdo al error generado por los neurones de la capa de salida.

Reconocimiento de Rostros

Como se mencionó previamente, el reconocimiento de rostros es una área dentro del reconocimiento de patrones que ha sido investigada desde hace varios años. Desde entonces se han propuesto varias técnicas para resolver esta tarea, entre ellas podemos mencionar las siguientes: las técnicas que utilizan igualamiento de plantillas[5], cálculo de *eigenfaces* [6], características geométricas [5], y los métodos que utilizan redes neuronales [6], entre otros. Para la implementación de este tipo de sistemas, independientemente de la técnica o metodología que sea implementada, se utilizan generalmente dos conjuntos de datos. El primer conjunto de datos es el que es utilizado durante la etapa de aprendizaje, el cual es comúnmente llamado *conjunto de entrenamiento*. Se debe de tratar que los patrones que integran este conjunto, sean lo más diferente posible entre sí, y que además, representen al problema, para poder obtener un buen porcentaje de generalización. El segundo conjunto de patrones, es el que se utiliza durante la etapa de reconocimiento y es llamado *conjunto de prueba*.

En la mayoría de estas técnicas, el reconocimiento de rostros involucra algunas de las siguientes etapas:

- preprocesamiento de la imagen,
- extracción de características,
- alguna forma de entrenamiento,
- y por último, la etapa de reconocimiento.

Preprocesamiento de la imagen

Dado que las imágenes de los rostros regularmente son tomadas en momentos diferentes, esto trae como consecuencia que las diferentes imágenes de una misma persona tengan variaciones en cuanto a la iluminación, la orientación, y el tamaño del rostro. Por tal motivo, es necesario que la imagen sea preprocesada antes de que pueda ser utilizada. Entre las tareas de preprocesamiento más comunes podemos encontrar las siguientes: extraer la imagen del rostro de una imagen más grande que contenga información irrelevante para el reconocimiento; normalización en cuanto al tamaño, es decir, que todas las imágenes de los rostros tengan un tamaño similar, y la aplicación de algún método de filtrado para mejorar la calidad de la imagen.

Extracción de Características

La extracción de características es una de las etapas de la cual depende en gran medida el buen desempeño del sistema de reconocimiento de rostros. El objetivo principal de esta etapa es extraer la información más discriminante de un rostro, eliminando la información que resulte irrelevante para el reconocimiento. En esta etapa podemos encontrar desde las técnicas que obtienen las características de manera manual, hasta técnicas más sofisticadas que extraen las características de forma automática. En varias técnicas, las principales características que son extraídas de los rostros son aquellas que describen a cada uno de los componentes básicos del rostro tales como los ojos, la nariz, la boca, la barba, las cejas, y por supuesto, la relación que existe entre ellos, es decir, la distancia entre cada uno de estos componentes.

Entrenamiento

El entrenamiento básicamente consiste en utilizar alguna forma de aprendizaje que le permita al sistema “aprender” los rostros que constituyen el conjunto de entrenamiento. El tipo de entrenamiento que sea utilizado para el aprendizaje, dependerá en gran medida, de la metodología que se esté utilizando para el reconocimiento. Por ejemplo, al utilizar una metodología como la de redes neuronales, el entrenamiento consiste en obtener los valores correspondientes a cada una de las conexiones (pesos) que constituyen a la red neuronal. Para este tipo de metodologías existen diferentes algoritmos para realizar el entrenamiento, entre los cuales el algoritmo de retro-propagación es uno de los más populares.

Reconocimiento

La última etapa que comúnmente es realizada en un sistema de reconocimiento de rostros, es la etapa de reconocimiento. Esta etapa básicamente consiste en alimentar al sistema con imágenes de personas diferentes a las utilizadas durante el entrenamiento, esperando obtener como resultado, alguna forma de codificación que nos permita identificar de que persona se trata.

Investigaciones actuales sobre reconocimiento de rostros

Enseguida se presentan algunos de los trabajos en el área de reconocimiento de rostros más relevantes de los últimos años. Roberto Brunelli y Tomaso Poggio [5], comparan dos estrategias sobre una base de datos común (imágenes frontales de los rostros de 47 personas; 26 masculinos y 21 femeninas, cuatro imágenes por persona). Ellos desarrollaron e implementaron dos nuevos algoritmos; el primero consiste en el cálculo de un conjunto de características geométricas, tales como ancho y alto de la nariz, posición de la boca y forma de la barba, y el segundo método está basado en el igualamiento de plantillas en niveles de gris. Ellos reportaron haber obtenido un 90% de reconocimiento usando características geométricas y un reconocimiento perfecto usando igualamiento de plantillas.

Meziane Yacoub y Younes Bennani [7] proponen un método para la selección de características para detectar e identificar zonas discriminatorias en la capa de entrada de un

modelo conexionista. El objetivo es seleccionar características simples que tengan un buen poder discriminatorio. El método que proponen es el ϵ HVS (HVS Extendido) que está basado en la medida HVS (Heurística para la Selección Variable) que ha mostrado resultado exitosos sobre Redes de Perceptrones Multicapa para la selección variable. Esta medida estima la contribución discriminatoria a la decisión final del clasificador de un peso dado. HVS permite identificar y seleccionar las variables más importantes en el espacio de características. Esto es logrado al eliminar las unidades redundantes y que tienen una baja contribución.

Ellos utilizaron un modelo conexionista convolucional llamado TDNN (de sus siglas en inglés Time Delay Neural Networks). Ellos utilizan una base de datos que consiste de 92 imágenes diferentes (4 personas diferentes, 23 imágenes por persona). La arquitectura del sistema que ellos utilizaron tiene dos módulos: el primero es un extractor de características con dos capas ocultas, con pesos compartidos, produciendo en la última capa oculta un mapa de características de la imagen presentada en la entrada, el segundo módulo es un clasificador que recibe como entrada el mapa de características obtenido anteriormente. La salida del clasificador es la identidad de la imagen presentada. La capa de salida tiene tantas unidades como personas se desee identificar. Ellos reportaron un porcentaje de reconocimiento del 100% sobre su base de datos.

Otro nuevo enfoque para el reconocimiento de rostros es el que proponen Oliver de Vel y Stefan Aeberhard [8]. Aquí ellos clasifican las técnicas para el reconocimiento de rostros en dos tipos, al primero lo llaman reconocimiento basado en características geométricas, el cual consiste en obtener un conjunto de características del rostro tales como ancho de la nariz, forma de la barba, espesor de las cejas, etc. Por otro lado, al segundo tipo lo llaman reconocimiento basado en imágenes, en donde mencionan que esta metodología de reconocimiento tienen una mayor simplicidad comparada con las metodologías basadas en características geométricas. Dentro de este tipo de metodologías podemos encontrar enfoques como los basados en plantillas y los basados en redes neuronales. Pero debido a que estos métodos, por lo general necesitan bastante tiempo para su entrenamiento y han sido probados con bases de datos de rostros relativamente pequeñas, ellos proponen una

tercera metodología, a la cual denominan: reconocimiento de rostros basado en líneas, la cual, tiene un buen porcentaje de reconocimiento, es tolerante a rotaciones, a variaciones en la escala y es invariante a la posición del rostro. Esta técnica consiste en representar a cada uno de los rostros como un conjunto de segmentos de líneas aleatorias obtenidas de la imagen de niveles de gris y utiliza un clasificador *nearest-neighbor*. Para la evaluación del desempeño de su sistema, utilizan la base de datos del “Olivetti & Oracle Research Laboratory” (ORL) y la base de datos del “University of Bern” (UB). Para realizar los experimentos, combinaron las dos bases de datos con el fin de obtener una base de datos más grande. Para la etapa de entrenamiento utilizaron 5 imágenes por persona y las otras 5 imágenes restantes las utilizaron para la etapa de reconocimiento. Como resultados de sus experimentos, reportan haber obtenido un porcentaje del 98% sobre la base de datos combinada, utilizando imágenes aleatorias para el entrenamiento y un porcentaje del 100%, utilizando un conjunto de imágenes seleccionadas para el entrenamiento.

