

Universidad Nacional de La Plata

Facultad de Informática

Técnicas biométricas, modelado del habla y sus aplicaciones.

Objetivo: Dar una visión general del estado del arte y de los ejes de investigación en el campo de la Biometría con especial énfasis en el área del tratamiento del habla.

El seminario estará dividido en tres partes principales que nos permitirán llevar al estudiante de manera progresiva a la asimilación de conceptos y técnicas en el campo de la biometría. La primera parte sitúa al estudiante dentro del área de estudio. La segunda le proporciona los conceptos de base a través de ejemplos concretos. La tercera presenta la complejidad del área.

1. Introducción

1.1 Aspectos generales y definición de Biometría.

1.2 Modalidades Biométricas.

1.2.1 Fisiológicas:

- Las huellas dactilares.
- La forma de la mano.
- La cara.
- El iris y la retina.
- La voz.

1.2.2 Del comportamiento:

- La voz.
- La escritura en el teclado.
- La forma de caminar.

1.2.3 Otras modalidades.

1.3 Contexto, aplicación y repercusión de los sistemas biométricos.

En esta primera parte el objetivo es dar un panorama general de la biometría. A través de un estudio general de ventajas, desventajas y de su ámbito de aplicación se presentará un conjunto representativo de modalidades biométricas. Se revisarán los elementos básicos que conforman un sistema biométrico con especial énfasis en los elementos físicos (rasgos morfológicos) y de comportamiento que se utilizan para caracterizar a una persona.

Con este primer acercamiento a la biometría el estudiante deberá ser capaz de distinguir los alcances, las limitaciones y las implicaciones de las técnicas básicas utilizadas en la biometría.

2 Verificación de locutor.

2.1 Diferencias entre el reconocimiento y la verificación de locutor.

2.2 Medidas de desempeño de un sistema de verificación.

2.3 Módulos de un sistema de verificación de locutor.

2.3.1 Caracterización de la señal de voz:

- producción,
- percepción y,
- normalización.

2.3.2 Modelado por:

- mezcla de Gaussianas (GMM),
- a base de Máquinas de Soporte Vectorial (SVM) y,
- por Modelos ocultos de Markov (HMM).

2.3.3 Normalización de resultados y la toma de decisión.

2.4 Campañas de evaluación.

El objetivo de esta parte del curso es que el estudiante adquiera los conocimientos básicos de las técnicas utilizadas en el reconocimiento de locutor. Esta presentación nos permitirá introducir conceptos generales a través del ejemplo: diferencia básica entre reconocimiento y verificación, reconocimiento dentro de un conjunto abierto o cerrado, medidas de desempeño de un sistema (curvas ROC y DET) y técnicas de adaptación de modelos. Al mismo tiempo nos dará la posibilidad de detallar el esquema general de los sistemas biométricos para el caso de

la verificación de locutor; por ejemplo: las características lingüísticas, como la lengua hablada, así como la dependencia o independencia del contenido del texto que transporta la señal de voz.

Las dificultades reales (bases de datos, condiciones de adquisición y de prueba) que suponen las técnicas presentadas se verán reflejadas en la presentación de un sistema de verificación de locutor y de las campañas internacionales para evaluarlo: evaluaciones NIST (National Institute of Standards and Technology) en reconocimiento de locutor.

Especial énfasis se hará en las implicaciones de la interpretación de los resultados desde el punto de vista de la aplicación.

Al final de esta segunda parte del seminario se busca que el estudiante tenga la capacidad de entender e implementar en su forma más básica los módulos que conforman un sistema de verificación de locutor.

3 Autenticación Multi-modal.

3.1 Independencia y complementariedad.

3.2 Fusión

3.2.1 Inter- e intra-modal.

3.2.2 De bajo y alto nivel.

3.2.3 Supervisada y no supervisada.

3.3 Modelos gráficos para la verificación de locutor.

3.3.1 Introducción a los modelos gráficos.

3.3.2 Fusión de características de la voz con Redes Bayesianas.

3.4 Bases de datos multi-modales.

En esta última parte el reconocimiento multi-modal será presentado como un área de investigación multi-disciplinaria en la que existen limitaciones debidas a su complejidad. Dadas las desventajas de cada modalidad y de las restricciones debidas a la aplicación en condiciones reales, la fusión de datos (modalidades) se presentará como una forma de mejorar el desempeño de los sistemas biométricos. El concepto de modalidad presentado en la primera parte se ampliará con la introducción de la intra-modalidad. Este concepto se empleará para describir las características que se obtienen de una misma fuente de información, por ejemplo: fusión de la imagen de la cara en 3D y en 2D o la fusión de la prosodia y las características espectrales de la señal de voz. Este ultimo ejemplo nos permitirá presentar los modelos gráficos y una de tantas de sus aplicaciones en la fusión de datos.

Como un ejemplo de la complejidad a la que se enfrentan los sistemas en condiciones reales de operación se presentarán los conceptos básicos y las diferencias existentes entre la fusión supervisada y no supervisada. De forma particular se abordará la fusión de alto nivel a través de pesos de ponderación. La dificultad de la sincronización inter-modalidad será presentada con la fusión de bajo nivel.

Al final de esta última parte el estudiante deberá haber tomado conciencia de los requerimientos y de las dificultades de un sistema real así como de los posibles ejes de investigación en el campo de la biometría.

Consultas:

Prosecretaría de Posgrado de la UNLP

Tel: 54-221-4236309

Mail: posgradounlp@presi.unlp.edu.ar

Página Web de la Facultad: www.info.unlp.edu.ar