

El algoritmo de dos fases para el reconocimiento de actividades en el contexto de la Industria 4.0 y los procesos dirigidos por humanos

Borja Bordel¹, Ramón Alcarria¹, Diego Sánchez-de-Rivera¹

¹Universidad Politécnica de Madrid,
Madrid, España

bbordel@dit.upm.es, ramon.alcarria@upm.es, diegosanchez@dit.upm.es

5 de octubre de 2022

Resumen. Los futuros sistemas de industria, la revolución conocida como la Industria 4.0, están considerando integrar gente en el mundo de Internet como potenciales prosumidores (entre proveedores de servicios y consumidores). En este contexto, el trabajo dirigido por humanos aparece como una realidad esencial y los instrumentos para generar bucles de información que se retroalimentan entre el subsistema social (personas) y el subsistema cibernético (componentes tecnológicas) son necesarios. A pesar de que muchos instrumentos distintos han sido propuestos, a día de hoy las técnicas de reconocimiento de patrones son las más prometedoras. Sin embargo, estas soluciones presentan algunos problemas pendientes. Por ejemplo, estos dependen del hardware seleccionado para obtener información de los usuarios; o presentan límites en el proceso de identificación. Para abordar esta situación, en este paper se propone un algoritmo de dos fases para integrar a las personas en los sistemas de la Industria 4.0. El algoritmo define acciones complejas como una composición de simples movimientos. Las acciones complejas son reconocidas utilizando los Movimientos Ocultos de Markov, y solo los movimientos simples se reconocen a través la DWT. De este modo, solo los movimientos dependen de los recursos de hardware empleados para capturar información, y se incrementa la precisión en el reconocimiento de acciones complejas. Una validación experimental será también llevada a cabo para evaluar y comparar la interpretación de los resultados propuestos.

Palabras clave: Industria 4.0; reconocimiento de patrones; DWT; Inteligencia Artificial; Modelos Ocultos de Markov.

1. Introducción

La Industria 4.0 [1] se basa en el uso sistemas Ciber-Físicos (uniones de procesos físicos y cibernéticos) [2] como componente tecnológica principal para futuras soluciones digitales, sobre todo (pero no únicamente) en el ámbito industrial. Por lo general, la digitalización ha causado, al final, el remplazo de los métodos tradicionales de trabajo por nuevos instrumentos digitales. Por ejemplo, los trabajadores en producción en cadena fueron sustituidos por robots durante la tercera revolución industrial.

Sin embargo, algunas aplicaciones industriales no pueden basarse en soluciones tecnológicas, por lo que la mano de obra se vuelve esencial [3]. Los productos hechos a mano son un ejemplo que reflejan la importancia del trabajo humano. Estos sectores de la industria, en cualquier caso, deben integrarse en la cuarta revolución industrial. Desde la unión de los Sistemas Ciber-Físicos (CPS) y el papel de las personas como proveedores de servicios (trabajos activos), surgen los CPS humanizados[4]. En estos nuevos sistemas, los procesos dirigidos por humanos están permitidos; los procesos[5] i.e. que son conocidos, ejecutados y controlados por gente (aunque serán supervisados a través de mecanismos digitales).

Para una verdadera integración entre la gente y la tecnología, que elimine los procesos de ejecución humana o tecnológica, para lo que se requieren técnicas de extracción de información. Durante los últimos años se han citado distintos métodos posibles, pero a día de hoy las técnicas de reconocimiento de patrones parecen las más prometedoras.

El uso de la IA, modelos estadísticos y recursos similares ha permitido un increíble desarrollo de los métodos de reconocimiento de patrones, aunque todavía quedan retos pendientes.

Inicialmente, las técnicas de reconocimiento de patrones dependen del hardware subyacente para la captura de información. La estructura y la memorización varía si (por ejemplo) empleamos sensores de infrarrojos en vez de acelerómetros. Es muy problemática la velocidad de evolución de las tecnologías hardware con respecto al software.

Por otro lado, existe un límite en la precisión del reconocimiento de procesos. De hecho, con las acciones humanas se vuelve más complicado, ya que se requieren más variables y modelos complejos para su reconocimiento. Esta aproximación genera grandes problemas de optimización con un error que aumenta al incrementar el número de variables, lo que provoca un decrecimiento del índice de reconocimiento[6]. En conclusión, las matemáticas (no hay software, por lo tanto, no depende de la implementación) aportan una cierta precisión al proceso de reconocimiento de patrones dando medidas para estudiar. Para evitar esta situación, deben ser consideradas menos variables. lo que reduce la complejidad del problema a la hora de analizarlo, una solución que no es aceptable en los ámbitos de la industria donde se desarrollan actividades de producción complejas.

Por lo tanto, el objetivo de este paper es describir un nuevo algoritmo de reconocimiento de patrones dirigido a estos dos problemas. El mecanismo propuesto define las acciones como una composición de movimientos simples. Estos se reconocen utilizando las técnicas de Deformación Dinámica del Tiempo (DTW) [7]. Este proceso depende del hardware empleado para la captura de información; pero las DTW son flexibles y la actualización del repositorio de reconocimiento es suficiente para configurar el algoritmo completo. Por otro lado, las acciones complejas son reconocidas como una combinación de movimientos simples a través de los Modelos Ocultos de Markov. HMM) [8]. Estos son totalmente independientes de la tecnología hardware , ya que solo consideran acciones simples. Estas dos fases se aproximan y reducen la complejidad de los modelos, incrementando la precisión y el éxito del índice en el reconocimiento de patrones.

El resto del paper se organiza como sigue: en la Sección 2 se describe el funcionamiento del reconocimiento de patrones para actividades humanas; en la Sección 3 se describe la solución propuesta, incluyendo las dos fases definidas; la Sección 4 presenta una prueba experimental utilizando un escenario real; y la en la sección se concluye el paper.

2. Estado del arte en el reconocimiento de patrones

Se han informado muchas técnicas diferentes de reconocimiento de patrones para actividades humanas. Sin embargo, la propuesta más común puede clasificarse en cinco categorías [9]: (i) Modelos ocultos de Markov; (ii) el campo aleatorio condicional de cadena de salto; (iii) Patrones Emergentes; (iv) el Campo Aleatorio Condicional; (v) clasificadores bayesianos

De hecho, la mayoría de los autores proponen el uso de Modelos Ocultos de Markov (HMM) para modelar las actividades humanas. HMM permite modelar acciones como cadenas de Markov. Básicamente, HMM genera estados ocultos a partir de datos observables. En particular, el objetivo final de esta técnica es construir la secuencia de estados ocultos que encaje con una determinada secuencia de datos. Para finalmente definir todo el modelo, HMM debe deducir de los datos los parámetros del modelo de manera confiable. La figura 1 muestra una representación esquemática de cómo funciona HMM. Cuando se reconocen las actividades humanas, las acciones que componen las actividades son los estados ocultos y las salidas de los sensores son los datos que se estudian. HMM, además, permite el uso de técnicas de entrenamiento considerando el conocimiento previo sobre el modelo. Este entrenamiento a veces es esencial para “inducir” todas las posibles secuencias de datos requeridas para calcular el HMM. Finalmente, es muy importante tener en cuenta que los HMM aislados simples se pueden combinar para crear modelos más grandes y complejos.

Los HMM, sin embargo, son inútiles para modelar ciertas actividades concurrentes, por lo que otros autores han reportado una nueva técnica denominada Conditional Random Field (CRF). Los CRF se emplean para modelar aquellas actividades que presentan acciones concurrentes o, en general, múltiples acciones que interactúan. Además, HMM requiere un gran esfuerzo de entrenamiento para descubrir todos los estados ocultos posibles. Para resolver estos problemas, el campo aleatorio condicional (CRF) emplea probabilidades condicionales en lugar de distribuciones de probabilidad conjunta. De esa forma, se pueden modelar fácilmente actividades cuyas acciones se desarrollan en cualquier orden. A diferencia de las cadenas en HMM, CRF emplea gráficos acíclicos y permite la integración de estados ocultos

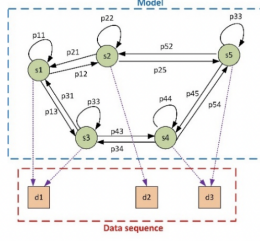


Figura 1: Representación gráfica de un HMM .

condicionales (estados que dependen de observaciones pasadas y/o futuras).

Los CRF, por otro lado, siguen siendo inútiles para modelar ciertos comportamientos, por lo que algunas propuestas generalizan este concepto y proponen el Skip Chain Conditional Random Field (SCCRF). SCCRF es una técnica de reconocimiento de patrones, más general que CRF, que permite modelar actividades que no son una secuencia de acciones en la naturaleza [14]. Esta técnica trata de capturar dependencias de largo alcance (cadena de salto); y puede entenderse como el producto de diferentes cadenas lineales. Sin embargo, al calcular este producto es bastante pesado y complicado, por lo que esta técnica suele ser demasiado costosa desde el punto de vista computacional para implementarla en pequeños sistemas integrados.

Otras propuestas emplean técnicas de descripción de mayor nivel como Emerging Patterns (EP). Para la mayoría de los autores, EP es una técnica que describe actividades como vectores de parámetros y sus valores correspondientes (ubicación, objeto, etc.). Utilizando distancias entre vectores es posible calcular y reconocer acciones desarrolladas por personas. Finalmente, otros autores han empleado con éxito técnicas secundarias como los clasificadores bayesianos, que identifican actividades haciendo una correspondencia entre las actividades humanas y las salidas más probables de los sensores mientras se realizan estas acciones, considerando que todos los sensores son independientes. Los árboles de decisión, las extensiones HMM y otras tecnologías similares también se han estudiado en la literatura, aunque estas propuestas son escasas.

Entre todas las tecnologías descritas, HMM no es la más poderosa. Sin embargo, encaja a la perfección con la Industria 4.0, donde las actuaciones son muy complejas pero muy estructuradas y ordenadas (según protocolos de empresa, políticas de eficiencia, etc.). Además, se requiere una retroalimentación rápida (a veces incluso en tiempo real) para garantizar que los procesos impulsados por humanos funcionen correctamente antes de que ocurra una falla crítica global. Por lo tanto, las soluciones computacionalmente costosas no son un enfoque válido, y estamos seleccionando HMM como tecnología de base principal. Para preservar su carácter liviano y, al mismo tiempo, poder modelar actividades complejas, introducimos un esquema de reconocimiento de dos fases que permite dividir acciones complejas en dos pasos más simples.

3. Algoritmo de reconocimiento de patrones en dos fases

Con el fin de (i) independizar el proceso de reconocimiento de patrones de usar aparatos de hardware para capturar información, (ii) permitir el reconocimiento de movimientos complejos, y (iii) preservar el carácter ligero de los modelos seleccionados, la solución propuesta una arquitectura con 3 capas tal que:

La capa más baja incluye la plataforma de hardware. Dispositivos de vigilancia como acelerómetros, teléfonos, sensores infrarrojos, etiquetas RFID, etc., se implementan para capturar información sobre el comportamiento de las personas. Las salidas de estos dispositivos crean secuencias de datos físicos cuyo formato, rango dinámico, etc., dependen totalmente de las tecnologías de hardware seleccionadas.

Estas secuencias de datos físicos luego se procesan en la capa intermedia usando técnicas de DTW. Como resultado, para cada secuencia de datos físicos, un simple movimiento o acción es reconocido. Estos movimientos simples se representan usando un formato de datos binarios para hacer la solución lo menos pesada posible. El software de este nivel debe modificarse cada vez que se actualiza la plataforma de hardware, pero las tecnologías DTW no requieren de un gran proceso de actualización, y actualizar el repositorio de patrones es suficiente para configurar el algoritmo en este nivel.

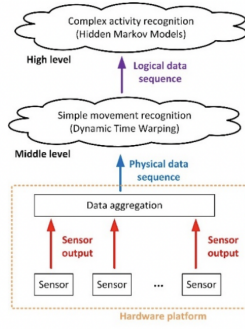


Figura 2: Arquitectura de la solución propuesta del patrón de reconocimiento.

Los movimientos simples reconocidos se agrupan para crear secuencias de datos lógicos. Estas secuencias alimentan un sistema de reconocimiento de patrones de alto nivel basado en los modelos ocultos de Markov. En este nivel, los componentes de software requieren un fuerte proceso de entrenamiento, pero la capa intermedia hace totalmente independiente la plataforma de hardware y de modelos de alto nivel. Por lo tanto, cualquier cambio en la plataforma de hardware no impone una actualización en el HMM, lo que sería extremadamente costoso computacionalmente. Por el análisis de la secuencia de movimientos simples, se reconocen acciones complejas.

La siguiente subsección describe en detalle las dos fases de reconocimiento de patrones propuestas.

3.1. Reconocimiento de movimientos simples: Dinámicas de Deformación del Tiempo

Para reconocer gestos o movimientos simples, se han seleccionado las técnicas de Deformación Dinámica del Tiempo. Estas tecnologías cumplen con los requisitos de los componentes de software de nivel medio, ya que se adaptan muy bien a las características de la plataforma de hardware fácilmente y son más rápidas y eficientes. (así los aparatos integrados las implementan).

En nuestra solución, el comportamiento humano se motriz a través de la familia de sensores símbolo S , conteniendo las N_s componentes (1).

$$S = \{ s_i, i = 1, \dots, N_s \} \quad (1)$$

Las salidas de estos sensores se muestran periódicamente como T_s : obteniendo para cada instante de tiempo, t , un vector de N_s valores (cada valor de cada sensor). Este vector Y_t se llama “muestra multidimensional simple” (2).

$$Y_t = \{ y_t^i, i = 1, \dots, N_s \} \quad (2)$$

Después, un movimiento simple Y tendrá una duración de T_m segundos y se describirá por la secuencia temporal de N_m muestras multidimensionales recogidas durante ese tiempo (3). Para el reconocimiento de movimientos tardíos, se crea un repositorio de patrón \mathcal{R} que contienen las secuencias temporales correspondientes para cada una de las acciones simples de K para reconocerlas (4).

$$Y = \{ Y_t, t = 1, \dots, T_m \} = \{ T^i, i = 1, \dots, N_m \} \quad (3)$$

$$\mathcal{R} = \{ R_i, i = 1, \dots, K \} \quad (4)$$

Por lo general, la gente realiza movimientos de un modo similar pero diferente. Por esto, las transiciones deben ser lentas o rápidas, algunas acciones elementales deben ser añadidas o eliminadas etc. Por lo tanto, dando una secuencia X N_x muestras, representando un movimiento para ser reconocido, deben estar localizadas en el patrón $R_i \in \mathcal{R}$ más cercano a X ; por lo que R_i se reconoce como una acción realizada. Para hacer esto se define la función distancia (5). Esta función que se aplica para calcular el valor de la matriz, necesita muestras que muchas veces no tienen la misma longitud ni están alineadas (6).

$$d : \mathcal{F} \times \mathcal{F} \longrightarrow \mathbb{R}^+, X^i, r_j^i \in \mathcal{F} \quad (5)$$

$$\mathcal{C} \in \mathbb{R}^{N_x \times N_m} \quad \mathcal{C}(n, m) = d(X^n, R_j^m) \quad (6)$$

En los sensores de posición (acelerómetros, aparatos de infrarrojos, etc) la función distancia es directamente aplicada a los sensores de salida (lo contrario ocurre, por ejemplo, en los micrófonos donde las salidas deben ser evaluadas en la carga del dominio). Aunque también pueden utilizarse otras funciones distancia (La divergencia simétrica de Kullback-Leibler o la distancia Manhattan), para este primer trabajo se utilizan las distancias Euclideas estándar (7).

$$d(X^n, R_j^m) = \sqrt{\sum_{i=1}^{N_s} (x_i^n - r_i^{mj})^2} \quad (7)$$

Después, se define la ruta de deformación $p = (p1, p2 \dots pL)$ como una secuencia de pares (nl, ml) con (nl, ml) símbolo pertenece $[1, N_x] \times [1, N_m]$ y $l \in [1, L]$, satisfaciendo las tres condiciones: (i) el límite, i.e. $p1 = [1, 1]$ y $pL = (F)$; (ii) la monotonía, i.e. (F); y (iii) la condición de intensidad de paso i.e. $p_l - p_{l-1}$ símbolo pertenece $(1, 0), (0, 1), (1, 1)$ con $\mathcal{L} \in [1, L-1]$.

Luego, el valor total de la ruta p_i se calcula añadiendo todos los valores o distancias parciales (8). Con todo esto, la distancia entre dos secuencias de datos R_i y X se define como el valor (distancia) de la ruta más óptima (9).

$$d_{pi}(X, R_j) = \sum_{\ell=1}^L \ell d(X_\ell^n, R_j^{m\ell}) \quad (8)$$

Finalmente, el movimiento simple es reconocido por la secuencia de datos X que es el que tiene el patrón R_i menor distancia (el más cercano) a X . Esta descripción es transigente con las variaciones de velocidad en la ejecución de movimientos, la introducción de nuevos microgestos, etc. Además, como se puede ver, cuando un hardware diferente se despliega, esto basta para actualizar el repositorio de la ruta \mathcal{R} para reconfigurar toda la solución al patrón de reconocimiento (ya que no se requiere entrenamiento).

3.2. Reconocimiento de movimientos complejos: Modelos Ocultos de Markov

El mecanismo propuesto anteriormente es muy útil para reconocer acciones simples, pero las actividades complejas involucran una gran cantidad de variables y requieren mucho más tiempo. Por lo tanto, DTW tiende a volverse impreciso y se requieren modelos probabilísticos. Entre todos los modelos existentes, HMM es el más adecuado para escenarios industriales y procesos humanos.

De la fase anterior, la infinidad de posibles movimientos simples de ser reconocidos es $\mathcal{M} = \{m_i, i = 1, \dots, K\}$. Además, se define un universo de estados $\mathcal{U} = \{u_i, i = 1, \dots, Q\}$, describiendo todos los estados que las personas pueden atravesar mientras interpretan cualquier acción bajo un estudio previo.

Entonces, una serie de observaciones $\mathcal{O} = \{o_i, i = 1, \dots, Z_o\}$ (movimientos simples reconocidos en la fase anterior) han sido consideradas, además de la secuencia de estados $\mathcal{V} = \{v_i, i = 1, \dots, Z_v\}$ describiendo la acción para ser modelada por HMM. En este caso inicial, asumimos cada observación correspondiente a un nuevo estado, así que $Z_o = Z_v$. Entonces, 3 matrices son calculadas: (i) la matriz transitoria A (10) describiendo la probabilidad del estado (uj) siguiendo el estado (ui); (ii) la matriz de observación (11) describiendo la probabilidad de observación (oj) causada por el estado (uj) independientemente de k; y (iii) la matriz de probabilidad inicial (12).

$$A = [a_{ij}] \quad a_{ij} = P(v_k = u_j \mid v_{k-1} = u_i) \quad (9)$$

$$B = [b_j(o_i)] \quad b_j(o_i) = P(x_k = o_i \mid v_k = u_j) \quad (10)$$

$$\Pi = [\pi_i] \quad \pi_i = P(v_1 = u_i) \quad (11)$$

Entonces, el HMM para cada actividad compleja(landai) para ser reconocida se describe gracias a estos 3 elementos previos(13).

$$\lambda_i = \left\{ A_i, B_j, \prod_i \right\} \quad (12)$$

Además, se hacen dos supuestos: (i) el de Markov(14) mostrando que cualquier estado depende solo del anterior; y (15) el de independencia, afirmando que cualquier secuencia de observación depende sólo del estado presente, no de estados previos u observaciones.

$$P(v_k | v_1, \dots, v_{k-1}) = P(v_k | v_{k-1}) \quad (13)$$

$$P(o_k | o_1, \dots, o_{k-1}, v_1, \dots, v_k) = P(o_k | v_k) \quad (14)$$

Para evaluar el modelo y reconocer la actividad que están realizando los usuarios, en este papel estamos utilizando un enfoque tradicional(16). Aunque los algoritmos directos han demostrado ser más eficientes, para este trabajo inicial estamos implementando directamente la expresión de evaluación en su forma tradicional.

16

El proceso de aprendizaje también se implementó en su forma más sencilla. Estadístico se emplearon definiciones para matriz transitoria, matriz de observación y matriz inicial. matriz de probabilidad. En particular, se empleó la definición de probabilidad de Laplace estimar estas tres matrices a partir de estadísticas sobre las actividades en estudio(17-19). El símbolo(\cdot) indica el número de veces que ocurre.

$$a_{ij} = P(u_j | u_i) = \frac{\text{contar}(u_j \text{ seguido de } u_i)}{\text{contar}(u_j)} \quad (15)$$

$$b_j(o_i) = P(o_i | u_j) = \frac{\text{contar}(o_i \text{ esta observado en el estado } u_j)}{\text{contar}(u_j)} \quad (16)$$

$$\pi_i = P(v_1 = u_i) = \frac{\text{contar}(v_1 = u_i)}{\text{contar}(v_1)} \quad (17)$$

4. Validación experimental: implementación y resultados

Con el fin de evaluar el cumplimiento de la solución propuesta, se ha designado una validación experimental y se ha llevado a cabo. Un escenario industrial ha sido imitado en algunas salas grandes de la Universidad Politécnica de Madrid. El escenario representó a una compañía tradicional de producción de productos hechos a mano. En particular, imitaron un pequeño PCB(placa de circuito impreso) productor.

Para conseguir información sobre el comportamiento de las personas, a varios participantes se les proporcionó guantes cibernéticos, incluyendo acelerómetros y un lector RFID. Los objetos cercanos al escenario fueron identificados con una etiqueta RFID, por lo que la plataforma de hardware propuesta puede identificar la posición de la mano y los objetos con los que las personas han interactuado.

Una lista de 12 actividades diferentes fue diseñada y definida usando la tecnología propuesta. La tabla 1 describe las 12 actividades definidas, además de una breve descripción sobre ellas.

18 personas estuvieron involucradas en el experimento. Las personas fueron solicitadas para realizar las actividades con números aleatorios. El orden verdadero, así como el de las actividades es reconocido y almacenado por un proceso de software supervisor. La tasa global de éxito para todas las soluciones ha sido evaluada, identificando, el mismo éxito para cada una de las distintas fases.

Para evaluar las mejoras obtenidas en comparación con las soluciones similares ya existentes, la misma secuencia de datos físicos ha sido empleada para alimentar la solución de un patrón estándar de reconocimiento basado solo en HMM. Usando un software de procesamiento de datos específicos, algunos resultados relevantes han sido obtenidos.

La figura 4 representa la tasa media de éxito para 3 casos: la solución global, la primera fase(DTW) y la segunda fase(HMM). Además, la tasa de éxito para el tradicional HMM-based acercamiento está incluido. Como se puede ver, la tecnología sugerida es, globalmente, alrededor de un 9% mejor que las técnicas tradicionales de reconocimiento de patrón basadas en HMM exclusivamente. Además,

Actividad	Descripción
Dibujar las rutas de los circuitos	El circuito para ser impreso es diseñado usando un programa PC de software específico
Imprimir el circuito diseñado usando un plóter	Usando fundas externas y un plóter, el circuito diseñado es impreso
Limpiar las caras del laminado de cobre	Usando un producto especial, todo el polvo y las partículas se eliminan de las caras del laminado de cobre
Copiar el circuito diseñado en las caras del cobre	El circuito diseñado en las fundas externas se copia en el laminado de cobre usando una explosión de rayos ultravioletas
Sumergir las placas en la piscina con ácido	Para eliminar lo innecesario del cobre, sumergimos las caras del laminado de cobre en un baño de ácido
Lavar el cobre con un baño de disolvente	Después del baño de ácido, la superficie del cobre restante se lava con un baño de disolvente
Capa de alineamiento	PCB están compuesto de varias capas, alineadas y amontonadas durante su fase
Inspección óptica	Se revisa la capa de alineamiento con un láser
Unir las capas exteriores con el sustrato	Usando un pegamento epóxico, las capas finales y exteriores de las placas se unen
Unir la placa	La unión ocurre en una mesa de acero pesado con abrazaderas metálicas
Taladrar los agujeros necesarios	Los agujeros para componentes, etcétera, son taladrados en el montón de placas
Enchapado	En un horno, la placa está lista

Cuadro 1: Tasa media de éxito para la solución propuesta.

la primera fase(basada en DTW) es alrededor de un 20% peor que la segunda fase(HMM) que es significativo que las DTW técnicas son más débiles por defecto.

5. Conclusiones y futuros trabajos

En este paper presentamos un nuevo patrón algorítmico de reconocimiento para integrar individuos en los sistemas de la Industria 4.0 y los procesos dirigidos por humano. El algoritmo define actividades complejas como una composición de movimientos simples. Las actividades complejas son reconocidas empleando los Modelos Ocultos de Markov, y los movimientos simples son reconocidos utilizando las DWT. Con el fin de activar el implemento de este algoritmo incrustado en objetos pequeños, se seleccionan configuraciones ligeras. Una prueba experimental de esto es llevarlo a cabo, y los resultados muestran una mejora global en un exitoso índice del 9

Los futuros trabajos considerarán metodologías más complejas para el procesamiento de datos, y la comparación para configuraciones distintas de la propuesta serán evaluadas. Además, la propuesta será analizada en diferentes escenarios.

6. Conclusiones y aportaciones de grupo

En el paper proporcionado se habla sobre la evolución de la industria y se hacen valoraciones de cómo influirán las nuevas tecnologías en la próxima revolución industrial, que dará paso a la Industria 4.0. Concretamente, se hace especial hincapié en la importancia de integrar el trabajo humano y el desarrollo tecnológico para implementar sector industrial. Por otro lado, se expone que el método más eficaz encontrado hasta la fecha sería el uso de reconocimiento de patrones en dos fases (DTW y HMM). A partir de esto se nos propone que extraigamos conclusiones sobre cómo podría llevarse a cabo el proceso. Para poder resolver el problema, nuestro grupo de trabajo se ha basado en un experimento donde se aborda el diseño de un reconocedor de voz. El sistema del reconocimiento de voz, consiste en capturar la voz para pasarla a través de una normalización, la cual al terminar tendrá

una transformación del tiempo a frecuencia, y que segmentaremos en fonemas para tener así ya lista la palabra a reconocer y determinar si es la palabra esperada. El experimento constó de 3 fases: Captura de voz Etapa de ajuste; el ajuste consiste en quitar la parte de silencio de la entrada a todos los archivos y dejarlos en un comienzo semejante para poder aplicar el algoritmo a señales con un comienzo en una palabra, sin que el silencio del inicio intervenga. Obtención de los coeficientes de Codificación de Predicción Lineal. Como habíamos visto en nuestro paper original, el reconocimiento de patrones se haría en 2 Fases: captura y transformación de acciones complejas y análisis de las acciones simples que obtenemos tras la fase 1. Para esto en el experimento se hicieron uso de 3 algoritmos distintos: Por un lado tenemos el algoritmo de ANN, que consta de 2 fases: la de aprendizaje y la de funcionamiento, y se basa en la formación de mapas topológicos. Durante la fase de aprendizaje se fijan los valores de las conexiones, lo que se conocería como red de neuronas. A continuación, se inicia un proceso de aprendizaje no supervisado y competitivo. Para esto se realizan diferentes pruebas donde enviamos diferentes señales de entrada. Como ya vimos, cada neurona está asociada a una información concreta. Así, cuando llega una señal de entrada (vector) con determinada información, las neuronas se activan, pero solo la que contiene la información correspondiente a la de nuestra señal de entrada consigue mantenerse activa. Por otro lado, durante cada ensayo con una misma señal la neurona modificará sus pesos para asemejarse más a la misma (se va formando la red de neuronas - mapa topológico) y así implementar el funcionamiento del sistema al permitir un mejor reconocimiento de una misma señal pese a las pequeñas fluctuaciones que pueda tener, pues estas quedan registradas. El algoritmo de aprendizaje autoorganizado basa su funcionamiento en las distancias euclídeas para determinar la neurona ganadora y la cual se asemejará a la entrada. Con todo esto se generaría un mapa que, en este caso, permitiría al sistema reconocer cada fonema con las diferentes variaciones en su pronunciación. Tras realizar todos los ensayos necesarios para la correcta configuración del sistema, el reconocimiento se basa en señales de entrada, cuando entra una palabra con los coeficientes del LPC, y la comparación de estas con los diferentes puntos del mapa.

El uso del algoritmo de DTW, tiene un funcionamiento similar, solo que en vez de emplearse los coeficientes de LPC se tomaría un patrón como base, ya que este está basado en el tiempo.

Y por último, el HMM se utilizó para la evaluación del algoritmo de Viterbi. Se inicia con una conversión de los LPC que toma los promedios obtenidos de cada una de las palabras de la base de datos registrados, comparando cada coeficiente de los promedios con cada coeficiente recibido de la pronunciación efectuada, para obtener una distancia euclídea, y así la menor distancia entre el coeficiente de la palabra recibida y los promedios es la letra a la que se convertirá dicha secuencia. Una vez tenida la secuencia de letras de la palabra pronunciada, se introduce al algoritmo de Viterbi para evaluar el comportamiento de la palabra y su probabilidad de que corresponda a una de las palabras guardadas en la base de datos.

En conclusión a esto podríamos decir lo siguiente. Por un lado, en el experimento quedó claro que es necesario ser muy riguroso a la hora de seleccionar los factores que van a influir en el proceso, así sea necesaria la buena elección en el de una base de datos, como una correcta determinación de las probabilidades del HMM para la evaluación del algoritmo de Viterbi. Finalmente, cabe añadir que, aunque esté demostrado que el DTW sea el mejor sistema para el reconocimiento, el empleo de un método de fases, donde queden involucrados nuevos métodos, permiten reforzar el sistema ya que así es posible compensar fallos, lo que sería imposible con el uso del algoritmo individual. En definitiva, la extrapolación de estos métodos al sector industrial nos permitiría obtener un sistema de trabajo mucho más eficaz.

Agradecimientos. La investigación que ha llevado a cabo los resultados obtenidos ha sido financiada por el Ministerio de Economía y Competitividad a través del proyecto SEMOLA (TEC2015-68284-R) y el Ministerio de Ciencias, Innovación y Universidades a través del proyecto VACADENA (RTC-2017-6031-2).

Referencias

1. BORDEL, B., ALCARRIA, R., SÁNCHEZ-DE-RIVERA, D., & ROBLES, T. (2017, NOVEMBER). PROTECTING INDUSTRY 4.0 SYSTEMS AGAINST THE MALICIOUS EFFECTS OF CYBER-PHYSICAL ATTACKS. IN INTERNATIONAL CONFERENCE ON UBIQUITOUS COMPUTING AND AMBIENT INTELLIGENCE (PP. 161- 171). SPRINGER, CHAM.
2. BORDEL, B., ALCARRIA, R., ROBLES, T., & MARTÍN, D. (2017). CYBER-PHYSICAL SYSTEMS: EXTENDING PERVASIVE SENSING FROM CONTROL THEORY TO THE INTERNET OF THINGS. *PERVASIVE AND MOBILE COMPUTING*, 40, 156-184.
3. NEFF, W. (2017). *WORK AND HUMAN BEHAVIOR*. ROUTLEDGE.
4. BORDEL, B., ALCARRIA, R., MARTÍN, D., ROBLES, T., & DE RIVERA, D. S. (2017). SELF-CONFIGURATION IN HUMANIZED CYBER-PHYSICAL SYSTEMS. *JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING*, 8(4), 485-496.
5. BORDEL, B., DE RIVERA, D. S., SÁNCHEZ-PICOT, Á., & ROBLES, T. (2016). PHYSICAL PROCESSES CONTROL IN INDUSTRY 4.0-BASED SYSTEMS: A FOCUS ON CYBER-PHYSICAL SYSTEMS. IN *UBIQUITOUS COMPUTING AND AMBIENT INTELLIGENCE* (PP. 257-262). SPRINGER, CHAM.
6. PAL, S. K., & WANG, P. P. (2017). *GENETIC ALGORITHMS FOR PATTERN RECOGNITION*. CRC PRESS.
7. MÜLLER, M. (2007). DYNAMIC TIME WARPING. *INFORMATION RETRIEVAL FOR MUSIC AND MOTION*, 69-84.
8. EDDY, S. R. (1996). HIDDEN MARKOV MODELS. *CURRENT OPINION IN STRUCTURAL BIOLOGY*, 6(3), 361-365.
9. KIM, E., HELAL, S., & COOK, D. (2010). HUMAN ACTIVITY RECOGNITION AND PATTERN DISCOVERY. *IEEE PERVASIVE COMPUTING/IEEE COMPUTER SOCIETY [AND] IEEE COMMUNICATIONS SOCIETY*, 9(1), 48.
10. LI, Z., WEI, Z., YUE, Y., WANG, H., JIA, W., BURKE, L. E., ... & SUN, M. (2015). AN ADAPTIVE HIDDEN MARKOV MODEL FOR ACTIVITY RECOGNITION BASED ON A WEARABLE MULTI-SENSOR DEVICE. *JOURNAL OF MEDICAL SYSTEMS*, 39(5), 57.
11. ORDONEZ, F. J., ENGLEBIENNE, G., DE TOLEDO, P., VAN KASTEREN, T., SANCHIS, A., & KROSE, B. (2014). IN-HOME ACTIVITY RECOGNITION: BAYESIAN INFERENCE FOR HIDDEN MARKOV MODELS. *IEEE PERVASIVE COMPUTING*, 13(3), 67-75.
12. ZHAN, K., FAUX, S., & RAMOS, F. (2015). MULTI-SCALE CONDITIONAL RANDOM FIELDS FOR FIRSTPERSON ACTIVITY RECOGNITION ON ELDERS AND DISABLED PATIENTS. *PERVASIVE AND MOBILE COMPUTING*, 16, 251-267.
13. LIU, A. A., NIE, W. Z., SU, Y. T., MA, L., HAO, T., & YANG, Z. X. (2015). COUPLED HIDDEN CONDITIONAL RANDOM FIELDS FOR RGB-D HUMAN ACTION RECOGNITION. *SIGNAL PROCESSING*, 112, 74-82.
14. LIU, J., HUANG, M., & ZHU, X. (2010, JULY). RECOGNIZING BIOMEDICAL NAMED ENTITIES USING SKIP-CHAIN CONDITIONAL RANDOM FIELDS. IN *PROCEEDINGS OF THE 2010 WORKSHOP ON BIOMEDICAL NATURAL LANGUAGE PROCESSING* (PP. 10-18). ASSOCIATION FOR COMPUTATIONAL LINGUISTICS.
15. GU, T., WU, Z., TAO, X., PUNG, H. K., & LU, J. (2009, MARCH). EPSICAR: AN EMERGING PATTERNS BASED APPROACH TO SEQUENTIAL, INTERLEAVED AND CONCURRENT ACTIVITY RECOGNITION. IN *PERVASIVE COMPUTING AND COMMUNICATIONS, 2009. PERCOM 2009. IEEE INTERNATIONAL CONFERENCE ON* (PP. 1-9). IEEE.
16. HU, B. G. (2014). WHAT ARE THE DIFFERENCES BETWEEN BAYESIAN CLASSIFIERS AND MUTUAL INFORMATION CLASSIFIERS?. *IEEE TRANS. NEURAL NETW. LEARNING SYST.*, 25(2), 249-264.
17. WANG, X., LIU, X., PEDRYCZ, W., & ZHANG, L. (2015). FUZZY RULE BASED DECISION TREES. *PATTERN RECOGNITION*, 48(1), 50-59.
18. DAVIS, M. H. (2018). *MARKOV MODELS & OPTIMIZATION*. ROUTLEDGE.
19. BORDEL SÁNCHEZ, B., ALCARRIA, R., MARTÍN, D., & ROBLES, T. (2015). TF4SM: A FRAMEWORK FOR DEVELOPING TRACEAB.