

Comunicación. Por el interior de HAPPY: Reconocimiento de Signos Gestuales QKQ

*María Montaña Morales Morgado. Juan Miguel León-Rojas
I.E.S. UNIVERSIDAD LABORAL. ESCUELA POLITÉCNICA, UEX*

INTRODUCCIÓN

HAPPY es un modelo de sistema computacional que ayuda durante el proceso de enseñanza-aprendizaje de varios sistemas de comunicación alternativa y aumentativa, principalmente de signos manuales y gráficos. Contiene *módulos generadores de tableros de comunicación en dos y tres dimensiones, predictores de letras, palabras y signos*, evaluadores de *habilidades motoras*, y herramientas para trabajar con *cuentos y juegos lingüísticos*.

Con HAPPY pretendemos (León-Rojas et al., 2000a y 2000b): (i) ayudar al aprendizaje de la Lengua de Signos Española (LSE), tanto a personas con necesidades especiales como a sus familiares, amigos, cuidadores, profesores, etc.; (ii) incluir otros sistemas de comunicación alternativa y aumentativa, en particular sistemas de signos gestuales y gráficos; (iii) servir de módulo en proyectos mayores, por ejemplo, en un proyecto de evaluación y rehabilitación, HAPPY aportaría información sobre las habilidades comunicativas y motoras en los sistemas de comunicación requeridos, además de ayudar al usuario en su aprendizaje; (iv) incluir un lenguaje donde esté representado cualquier dato de naturaleza vocal y no vocal; (v) romper la barrera de la distancia para la comunicación con signos, sean gestuales o gráficos, impulsando, por ejemplo, la comunicación con signos por Internet.

La mayoría de los signos de la LSE responden al que llamamos *patrón QKQ*: se reconocen a partir de dos *queiremas* (configuraciones manuales), uno “de salida” y el otro “de llegada”, y una trayectoria (*kinema*) donde se produce la transformación del primero en el segundo. Asumiendo el contexto (lenguaje controlado) de signos QKQ, entonces cualquier signo (QKQ) es isomorfo a una tríada (queirema, kinema, queirema), y por ende, reconocer cualquier signo (QKQ) equivale a reconocer una tríada (queirema, kinema, queirema).

Este trabajo pretende divulgar la metodología usada por HAPPY para el reconocimiento de signos QKQ.

RECONOCIMIENTO DE UN QUEIREMA

El reconocimiento de un queirema se lleva a cabo en base a su modelización *fuzzy* de tipo 2, a ciertas sub-relaciones de máxima similitud entre distribuciones de probabilidad, y a un proceso de inferencia bayesiana (León-Rojas et al., 2001). Como atributos de cualquier queirema tomamos el conjunto de ángulos de flexión y aducción de los dedos, y el arco palmar, que se supondrán por tanto conocidos en cual-

quier instante de muestreo. Estos valores angulares se normalizan en [0,1]. Cada queirema es un subconjunto *fuzzy* de tipo 2 del conjunto de ángulos, usando el conjunto de términos $L([0,1])=\{<flexionado> (F), <algo flexionado> (AF), <ni flexionado ni extendido> (NFE), <algo extendido> (AE), <extendido> (E), <pegado> (P), <algo pegado> (AP), <ni pegado ni separado> (NPS), <algo separado> (AS), <separado> (S), <irrelevante> (IRR)\}$. Por ejemplo, fijémonos en la descripción del queirema <LV> del LSE (Rodríguez, 1992) como conjunto *fuzzy* de tipo 2:

extensión/flexión		aducción/abducción				
		- I	- M	- A	- Ñ	Arco palmar
		B	B	RR	RR	E

SEGMENTACIÓN DE SIGNOS QKQ

Por segmentar un signo entendemos el hecho de detectar los instantes de tiempo inicial y final de su articulación. En este proceso de segmentación interviene una *jerarquía de agentes computacionales*. Desde el grupo de recogida de datos angulares, meros agentes recolectores de la secuencia cronológica de los datos angulares preprocesados, hasta los agentes encargados de reconocer específicamente cada componente articulatoria del signo.

Detectar movimientos significa reconocer cambios de coordenadas espaciales en el tiempo. El reconocimiento de la dirección de movimiento (*kineprosema*) es inmediato a partir del conocimiento de las variaciones de posición (*toponemas*) con respecto al tiempo. La identificación de los tipos de movimiento de la mano, considerada en su totalidad, es tan sencilla (o tan compleja) como el reconocimiento de una *trayectoria determinada en el espacio 3D*. Dicha trayectoria se aproxima por una poligonal, muestreando con una frecuencia de 10 tomas por segundo. Según nuestra experiencia, la característica más discriminante es el producto vectorial relativo a cada par de vectores consecutivos (Fig. 1).

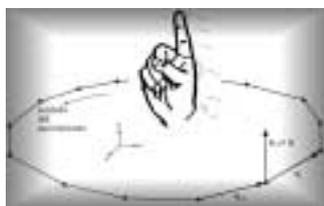


Fig. 1. Producto vectorial relativo

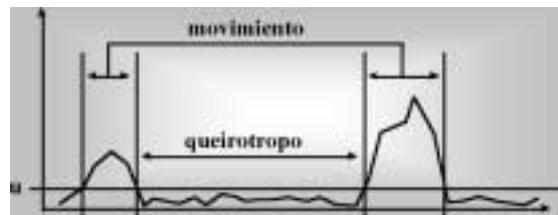


Fig. 2. El eje de ordenadas representa el rango de S; el de abscisas es un eje de un grupo unificador.

La metodología consiste en que ciertos *agentes segmentadores* evalúen un conjunto de activadores o indicadores de movimiento, conjunto actualmente en estudio, de tal forma que la importancia relativa de cada uno de sus miembros puede ponderarse según determinados factores de relevancia. La suma S, de las valoraciones totales de estos activadores, se compara con un valor umbral u, de presencia de movimiento (véase la Fig. 2). Si $S < u$, se estima que no hay movimiento. Así se segmenta un *queirotopo* = queirema + *queirotopema* (orientación) + localización relativa (toponema). En la zona donde se ha detectado movimiento habrá de reconocerse el tipo de movimiento (*kinema*) (en línea recta, en arco, bucles, etc.)

Existe un grupo *unificador*, encargado de realizar, siempre que sea posible, la unificación de ciertas componentes de registros consecutivos. Por ejemplo, imaginemos una secuencia de 10 registros, donde la información del *agente inspector de movimiento* ha sido la que muestra la Fig. 3. El *agente reconocedor de kinemas* es el encargado de identificar, en este caso, un zig-zag. También debe interpretar ciertos caminos poligonales, por ejemplo $[t_i, t_j]$, como líneas rectas (aquí las pendientes consecutivas no presentan diferencias superiores a un umbral prefijado). En general, cualquier información aportada por cualquier agente, en un registro, puede ser similar a la del registro siguiente, o tener un atisbo de continuidad. Por ello, a cada agente de los grupos anteriores le corresponde un *agente unificador*.

ADQUISICIÓN DE DATOS MEDIANTE SISTEMAS MEDIDORES DE POTENCIALES ELÉCTRICOS

La metodología expuesta es independiente del modo de adquirir los datos. Tres podrían ser las posibilidades: *cámaras de video*, *guantes con sensores* y *sistemas medidores de potenciales eléctricos* (SMPE). Esta última es la más atractiva. Los SMPE se basan en electrodos de superficie captadores de diferencias de potencial asociadas a la actividad muscular (electromiográfica –EMG–) y a las ondas (alfa) emitidas por el cerebro (actividad electroencefalográfica –EEG–). Teniendo muy presente la portabilidad del sistema y el deber paliar o mejor, enteramente anular un posible sentimiento de robotización en el usuario, se piensa en este tipo de interfaces. La propuesta de Hiraiwa et al. (1992), se basa en 2-4 electrodos de superficie que, colocados en el antebrazo miden las señales EMG. La implementación se basa en una arquitectura de red neuronal, cuyo entrenamiento se realiza mediante un VPL DataGlove, que permite identificar de forma supervisada los valores EMG correspondientes a cada gesto. El sistema se integra en una pulsera.

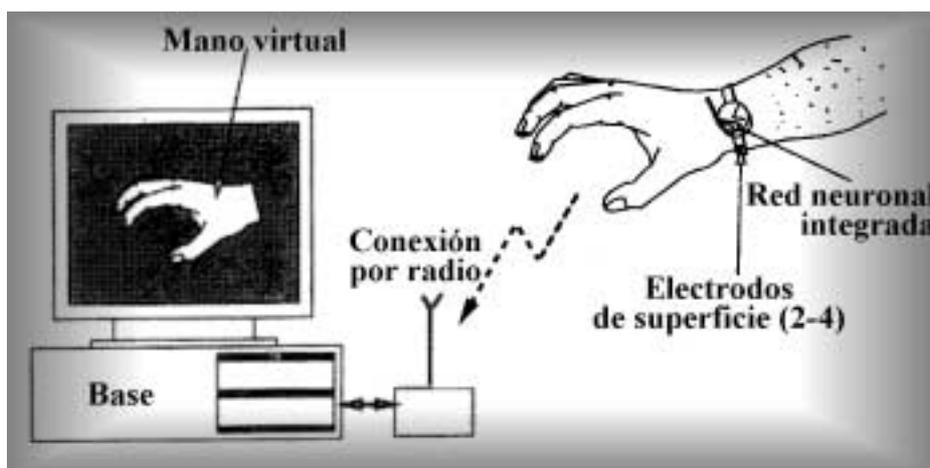


Fig. 4. Sensor de EMG y red neuronal, integrados en un reloj de pulsera. Al realizarse la conexión por radio se elimina la unión física (cableado) del usuario (receptor) con el emisor.

EXPERIMENTACIÓN

Debido a la carencia de un número suficiente de expertos signantes, se ha simulado. Adoptamos el *modelo anatómico computacional* de Dorner (1994) –basado en Buchholz y Armstrong (1992)–, que proporciona rangos interdependientes para todos los ángulos. Los términos de $L([0,1])$ son números fuzzy triangulares. Las *tasas de reconocimiento* de signos QKQ, tras un entrenamiento con 200 expertos y 50 articulaciones (10.000 patrones por signo), han sido del 91% para *expertos registrados*, del 86% para

usuarios registrados y del 70% para *usuarios no registrados*. Resulta ser robusto: proporciona los mismos resultados hasta con un 6% de ruido añadido a cualquiera de los ángulos y/o posiciones relativas.

AGRADECIMIENTOS

Agradecemos a la Junta de Extremadura, su ayuda original a través del PRI-97-C1063.

REFERENCIAS

BUCHHOLZ, B. y AMSTRONG, T. J. (1992) "A kinematic model of the human hand to evaluate its prehensile capabilities". *Journal of Biomechanics*, 25(2): 149-162.

DORNER, B. (1994). *Chasing the Colour Glove: Visual Hand Tracking*. Simon Fraser University. (MS. Thesis, junio).

HIRAIWA, A., UCHIDA, N., SONEHARA, N. y SHIMOHARA, K. (1992) "EMG pattern recognition by neural networks for prosthetic fingers control "CyberFinger", *Proceedings of the Second International Symposium on Measurement and Control in Robotics (ISMCR'92)*, Tsukuba.

LEÓN-ROJAS, J. M., MORALES, M., MORENO, J., SILVA, A. y MASERO, V. (2000) "HAPPY: Un modelo de sistema de ayuda por computador para el aprendizaje de sistemas de comunicación basados en signos gestuales y gráficos". *Tercer Congreso Iberoamericano de Comunicación Alternativa y Aumentativa*. Instituto de Automática Industrial (CSIC) - IMSERSO - CYTED, pp. 41-44. Madrid.

LEÓN-ROJAS, J. M., MORENO, J., SILVA, A., MASERO, V., MORALES, M., BARRIGÓN, J. M. (2000) "E/A asistida por computador de sistemas de comunicación basados en signos", *VII Congreso de Innovación Educativa en Enseñanzas Técnicas*, San Sebastián.

LEÓN-ROJAS, J. M., MASERO, V. y MORALES, M. (2001) "Inferencia bayesiana de queiremas usando conjuntos *fuzzy* de tipo dos". *XXVI Congreso Nacional de Estadística e Investigación Operativa*. Universidad de Jaén. Úbeda, 6-8 de noviembre.

RODRÍGUEZ, M. A. (1992) *Lenguaje de Signos*. Confederación Nacional de Sordos de España - Fundación ONCE, Madrid.