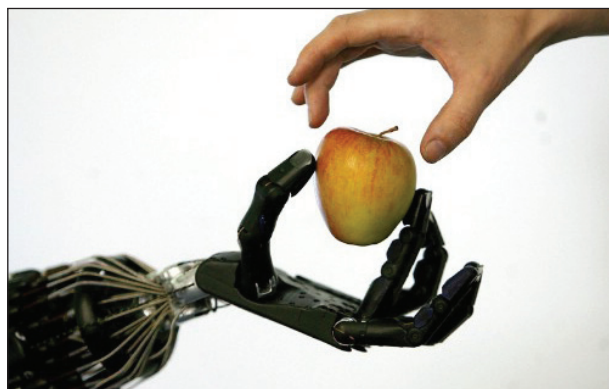


Imitación de gestos por brazos robóticos: Una propuesta para evaluar su calidad



Sandra Nope-Rodríguez * Doctora en Ingeniería
Humberto Loaiza-Correa * Doctor en Robótica
Eduardo Caicedo-Bravo * Doctor en Informática Industrial

* Universidad del Valle. Grupo de Percepción y Sistemas Inteligentes,
Ciudad Universitaria Meléndez, Calle 13 No 100-00 – Cali (Colombia). Tfno: +57 2
3303436. sandra.nope@correounivalle.edu.co, humberto.loaiza@correounivalle.edu.co,
eduardo.caicedo@correounivalle.edu.co

Recibido: 02/02/2010 • Aceptado: 03/03/2010

Imitation of gestures by robotic arms: A quality assessment approach

ABSTRACT

- This article proposes quantitative evaluation metrics of imitation, suited for the imitation of four gestures performed by a robotic arm on a simulation environment; for each gesture, the metric relies on a group of simple geometric shapes. This approach broadens the concept of imitation; since common metrics just accumulate the error between the joints of the demonstrator and the imitator along their trajectories, in other words, check for the exact replica of the gesture performed by the demonstrator. The results of the proposed metric are compared with an evaluation survey conducted in human observers.
- **Key words:** Learning by demonstration, imitation, metrics, robotic.

RESUMEN

En este artículo se aborda una propuesta de evaluación cuantitativa de la imitación de cuatro gestos realizados por un brazo robótico simulado; para cada uno se especifica una métrica simple basada en formas geométricas. Este enfoque permite ampliar el concepto de imitación, ya que hasta ahora las métricas estiman el error entre la trayectoria seguida por las articulaciones del demostrador respecto a las del imitador, en otras palabras, verifican que se realice una réplica exacta del gesto realizado por el demostrador. Los resultados de las métricas propuestas se comparan con los de una encuesta realizada a observadores humanos.

Palabras clave: aprendizaje por demostración, imitación, métricas, robótica.

1. INTRODUCCIÓN

La imitación es la forma natural en la que muchos seres vivos aprenden. En la actualidad, se han logrado avances en la programación de robots a través de la imitación, también denominada aprendizaje por demostración, especialmente con la evolución de los robots humanoides y otros sistemas altamente articulados [1].

Entre las ventajas más destacables de esta programación, se encuentran la facilidad con la que se podrían programar robots simultáneamente para que realicen tareas complejas, y la posibilidad de que los robots difieran morfológicamente del demostrador o maestro (que puede ser un humano u otro robot).

Sin embargo, la fidelidad o precisión en la reproducción puede variar y no existe ningún método que proponga niveles, límites o umbrales para definir el éxito en la imitación, aparte de que puede haber fallos en la asociación entre los comportamientos observados.

En este sentido, el observador juega un papel crucial, ya que debe juzgar o evaluar la calidad de la imitación. El observador puede ser el demostrador, el imitador, o un ente externo (humano ó máquina). El observador, por tanto, puede atribuir un amplio rango de objetivos y/o metas diferentes al imitador y/o al demostrador para juzgar la similitud entre las acciones.

Debido a la buena capacidad humana para imitar y reconocer correspondencias entre acciones, se han utilizado encuestas para determinar la calidad de la imitación de los robots, como se expone en [2].

A pesar de la subjetividad involucrada a la hora de juzgar si hay imitación o

no, es de vital importancia cuantificar el desempeño del robot durante la ejecución de la imitación, comprobando la similitud entre la acción del modelo y la del imitador.

Las mediciones cuantitativas de evaluación - hasta ahora - suelen verificar el error entre la posición de un objeto o de las articulaciones del robot respecto a las del demostrador, por ejemplo en [3]. Sin embargo, usando un criterio más amplio de lo que es imitación, este “error” no necesariamente es el más apropiado, especialmente si demostrador e imitador no comparten las mismas características morfológicas.

Las métricas asociadas a la imitación, tal como se ha reconocido [4], están directamente relacionadas con el diverso rango de aplicaciones humano-robot y del concepto de imitación usado. A pesar de su importancia, son escasos los trabajos sobre medidas de similitud en la imitación no exacta de la acción ejecutada por un demostrador, tal como se reconoce en [5]. Para paliar esta deficiencia, se proponen cuatro métricas que permiten cuantificar la imitación de cuatro gestos realizada por un brazo robótico simulado ante diferentes demostradores humanos. Estas métricas son bastante sencillas y se fundamentan en trabajos de neurociencia, ciencias cognitivas y psicología que indican que las personas fijan su atención en la trayectoria seguida por el efector final [6]. Así, para cada gesto utilizado, se asoció una figura geométrica a la forma de la trayectoria durante su ejecución, debido a que lo que se pretende es evaluar la calidad de la imitación de un gesto, y no el reconocimiento y clasificación de gestos. Esta propuesta de asociar figuras geométricas a las formas de las trayectorias, es simple, robusta y extensible a otra clase de gestos.

En el apartado 2 se describe el sistema que se validó mediante las métricas descritas y las condiciones experimentales. En el apartado 3 se explican las ideas de cada una de las métricas y su desarrollo matemático y computacional. Los resultados y su análisis correspondiente aparecen en el apartado 4; y las conclusiones y trabajo futuro en el 5.

2. DESCRIPCIÓN DEL SISTEMA

Se describe los principales bloques del sistema, de acuerdo con la definición de imitación dada por [7]: “cuando un gesto es observado, reconocido y ejecutado” (ver Fig.1).

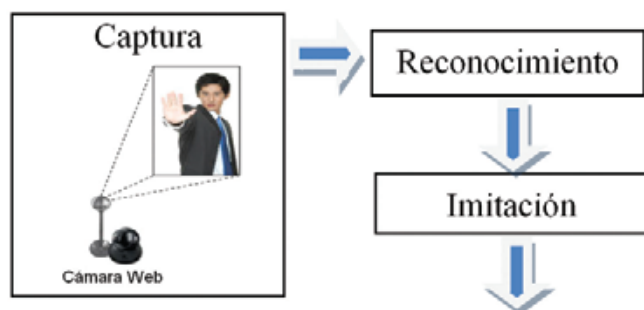


Fig. 1: Diagrama de Bloques General del Sistema

El bloque de “Captura” extrae la información visual (vídeos) de la escena mediante una cámara Web. La captura de los vídeos se inicia y termina manualmente, así como el enfoque al demostrador.

El bloque de “Reconocimiento” recibe como entrada los vídeos tomados por el bloque precedente, y genera a partir de ellos una representación bio-inspirada [8] que reúne la información del movimiento, y que presenta alta discriminación para los gestos escogidos.

El bloque de “Imitación” relaciona la representación de la información con el espacio motor. De esta manera, se dota al sistema de una transformación directa a variables articulares del robot, que permiten que el robot realice su propia interpretación de la escena observada. En [9], se describe la manera de obtener los valores articulares desde una imagen monocular tomada por una cámara Web sin calibrar, de tal forma que un brazo robótico con 6 grados de libertad (*Degrees Of Freedom* - DOF) pueda imitar la pose de un brazo humano.

La transformación se obtuvo entrenando una *Red Neuronal de Regresión Generalizada* (RNRG) con los datos obtenidos anteriormente. El modelo neuronal de regresión generalizada fue propuesto y desarrollado por **Specht** en 1991 (Referenciado [10]), y se basa en la teoría de regresión no lineal, y es, en esencia, un método que permite estimar una función de densidad conjunta a través de un conjunto de entrenamiento. Esta red permite lograr buenas interpolaciones, pudiendo aproximar cualquier función arbitraria entre vectores de entrada y salida directamente a partir de los datos de entrenamiento. Además, no requiere entrenamiento iterativo y basta con que se explore el efecto de su único parámetro a definir, llamado parámetro de suavizado.

1.1. CONDICIONES EXPERIMENTALES

Las condiciones de grabación de los vídeos en formato .AVI fueron semi-controladas. Se utilizó iluminación de lámparas tipo oficina en horas diurnas, una cámara Web QuickCam Logitech Orbit^{MT}, con imágenes con una resolución de 240x320 píxeles², con una velocidad de captura de 30 tramas/sg. Se evitó el solapamiento del objeto de interés (extremidad superior) con objetos con características similares de color. El procesamiento de los vídeos se realizó en Matlab.

1.2. CRITERIOS DE EVALUACIÓN

Para evaluar la funcionalidad del sistema se empleó un grupo de cuatro gestos que fueron observados, reconocidos e imitados. Los gestos corresponden a:

- Gesto 1: rotar la mano en sentido contrario a las manecillas del reloj y devolverse, como cuando se saluda.
- Gesto 2: bajar y subir la mano, como cuando se llama a una persona.
- Gesto 3: rotar la mano en sentido inverso a las manecillas del reloj.

- Gesto 4: acercar y alejar la mano respecto a la cámara.

Los tres primeros gestos fueron empleados en el trabajo de [11], ya que estos gestos, en general, sólo requieren el movimiento de 2 de los 6 DOF disponibles en las simulaciones del brazo robótico, se adicionó el gesto 4 que requiere del movimiento simultáneo de todas las articulaciones.

Este trabajo corresponde a una primera exploración experimental, donde se grabaron 140 vídeos, con diferentes demostradores realizando uno de los cuatro gestos posibles, es decir, 35 interpretaciones (vídeos) de cada gesto. De éstos, 48 (12 de cada gesto) fueron usados en la fase inicial de entrenamiento de la RNRG y 92 (23 de cada gesto) para evaluación del sistema.

El criterio para validar los resultados de la evaluación de la imitación obtenida mediante las métricas propuestas, consiste en encuestar a un grupo de personas (ocho), para que determinen si, a su criterio, las simulaciones del brazo robótico corresponden a las acciones realizadas por el demostrador humano; y contrastar estos resultados con los obtenidos a través de una evaluación cuantitativa. Esto se hace necesario debido a la inexistencia de trabajos que evalúen la imitación de los mismos cuatro gestos empleados. Adicionalmente, debido a que las líneas de investigación en imitación robótica, apuntan a dos direcciones: una propende por una réplica exacta de la acción realizada por el demostrador (mímica), que suelen utilizar métodos cuantitativos para comparar trayectorias o valores articulares; y otra que tiene como propósito que el robot realice su propia interpretación de los gestos observados, y que suelen utilizar métodos cualitativos de evaluación a través de encuestas. Dado que en este trabajo se pretende cuantificar la imitación de un robot a partir de su interpretación de la acción observada, se emplean los dos mecanismos de evaluación para validar los resultados provistos por las métricas propuestas.

3. MÉTRICAS PARA EVALUAR LA IMITACIÓN

Las métricas para evaluar la imitación suelen comparar las trayectorias del demostrador con las del imitador. Esto es, claramente, poco apropiado cuando se pretende que el imitador realice su propia interpretación; además, porque en el mundo real una réplica exacta o mímica, no es necesariamente la respuesta óptima. Esto se debe, entre otras cosas, a que el demostrador y el imitador no comparten las mismas medidas antropomórficas y condiciones de tarea.

Los indicadores propuestos se obtuvieron analizando la trayectoria del dedo corazón de la mano (efector final) del demostrador. Esta selección permite al imitador concentrarse en lo más destacado del movimiento y permite atenuar la alta variabilidad en los comportamientos motores de los humanos que usan para realizar una tarea.

Por otro lado, la observación de la trayectoria seguida por el dedo corazón, durante la ejecución realizada por diferentes demostradores sugiere que éstas tienen una forma

determinada para cada uno de dichos gestos, es especialmente notable en los gestos del 1 al 3. A continuación se describe como se obtiene la trayectoria 2D desde un vídeo, que posteriormente es usada en las métricas.

3.1. PROYECCIÓN DE COORDENADAS

El primer paso consiste en proyectar los movimientos 3D del brazo robótico en un plano de imagen cualquiera. Para ello, se parte de los valores de las variables articulares que entrega el bloque “Imitación” para cada DOF del brazo, y se utilizan para estimar las coordenadas 3D del efector final (X , Y , Z). Estas coordenadas se transforman en 2D siguiendo un modelo ortográfico, de acuerdo con la ecuación (1).

$$\begin{pmatrix} u \\ v \end{pmatrix} = s \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (1)$$

El valor de la constante s puede ser cualquiera, lo que verifica la independencia a la escala de las métricas propuestas; el valor utilizado fue de $s=10$, que se empleó durante la reconstrucción tridimensional necesaria para el entrenamiento del sistema.

Una vez transformados los movimientos del brazo a coordenadas 2D, es posible cuantificar la calidad de la imitación de cada gesto tal como se explica a continuación.

3.2. MÉTRICA PARA EL GESTO DE SALUDAR

La forma de la trayectoria seguida por el dedo corazón durante la ejecución del gesto de saludar, corresponde aproximadamente a dos parábolas, en donde la variable dependiente corresponde al eje de las abscisas; una de las parábolas corresponde al recorrido de la mano durante la rotación en el sentido inverso a las manecillas del reloj (desde la perspectiva del observador) y la otra al recorrido durante el regreso de la mano. Aunque se habla de parábolas diferentes, pueden coincidir.

La trayectoria se divide en dos: del inicio en la ejecución hasta el menor valor en las coordenadas de las abscisas, y de ahí hasta el final de la ejecución. Así, para cada subgrupo se estiman los mejores valores las constantes a , b y c que define la parábola en la ecuación (2), empleando el método óptimo de Ransac (*RANdom Sample Consensus*).

$$u = av^2 + bv + c \quad (2)$$

Para la estimación de los parámetros óptimos, se eligen al azar el 2% de los pares coordenados de la sub-trayectoria y se calculan los valores de a , b y c , y se evalúa el error producido con estos valores para todos los puntos dándoles un margen de error (0.5 píxeles para este caso). Este procedimiento se repite 1000 veces y, al final, se escoge el valor de las constantes con las que se logró el mejor ajuste a todos los pares coordenados. Este número de puntos que se ajustan, define el porcentaje de éxito en la ejecución del gesto 1.

La Fig.2 muestra, en azul, la trayectoria seguida por la marca sobre el dedo corazón durante la ejecución del gesto 1 por parte de un demostrador. Los asteriscos fucsias (*) corresponden a la mejor parábola que encontró el sistema que se ajusta a los pares coordenados durante el inicio del gesto 1, y las cruces verdes (+) en segunda parte del gesto. La figura en el recuadro corresponde a una representación en el plano de la imagen, y el fondo corresponde a una ampliación de la trayectoria. En el recuadro los datos parecen corresponder apropiadamente a una parábola, pero en la ampliación se aprecian un grupo de puntos al final de la ejecución, que no coinciden exactamente con la forma de una parábola.

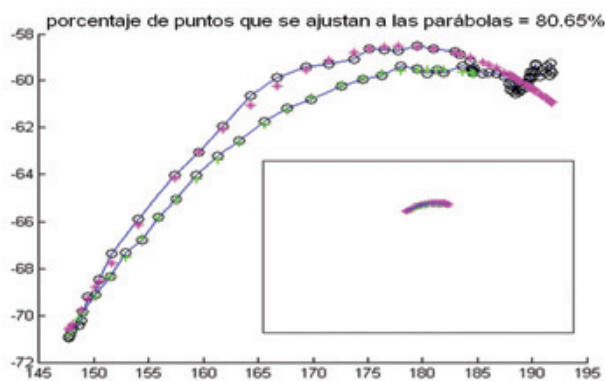


Figura 2: Ejemplo de la métrica usada para el gesto 1

En ambas figuras, los asteriscos fucsias (*) corresponden a la mejor parábola encontrada por el sistema que se ajusta a los pares coordenados durante el inicio del gesto 1 (círculos azules), y las cruces verdes (+) a la mejor parábola que se ajusta a los pares coordenados durante la ejecución de la segunda parte del gesto.

3.3. MÉTRICA PARA EL GESTO DE ABANICAR

En el gesto 2 (bajar y subir la mano - abanicar), la forma corresponde aproximadamente a dos parábolas pero, en este caso, la variable dependiente corresponde al eje de las ordenadas, por lo que el modelo se define de acuerdo con la ecuación (3). Al igual que en el caso anterior, las parábolas pueden coincidir y las constantes a , b y c que determinan las parábolas se estiman mediante el método de Ransac.

$$v = au^2 + bu + c \quad (3)$$

La Fig.3 muestra, en azul, la trayectoria seguida por la marca sobre el dedo corazón durante la ejecución del gesto 2 por parte de un demostrador. Los asteriscos fucsias (*) corresponden a la mejor parábola que encontró el sistema que se ajusta a los pares coordenados durante el inicio del gesto 2, y las cruces verdes (+) en segunda parte del gesto. La figura en el recuadro corresponde a una representación en el plano de la imagen, y el fondo corresponde a una ampliación de la trayectoria.

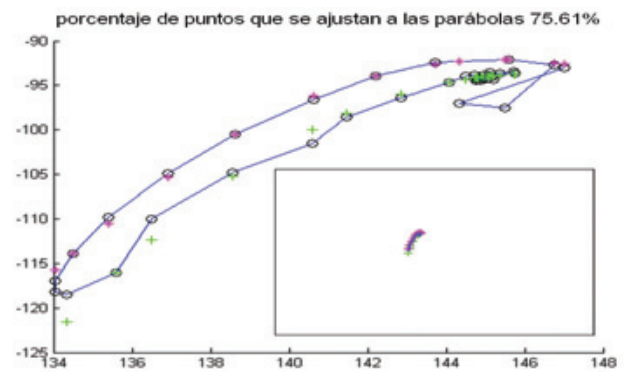


Fig. 3: Ejemplo de la métrica usada para el gesto 2

Aunque se podría pensar en realizar la estimación de las parábolas para el gesto 2, usando el procedimiento del gesto 1, se verificó que se obtiene un mejor ajuste de los puntos utilizando las dos aproximaciones planteadas para cada gesto.

3.4. MÉTRICA PARA EL GESTO DE ROTAR LA MANO

En el gesto 3 (rotar la mano), la forma de la trayectoria corresponde aproximadamente a una elipse. Sin embargo, ésta no está centrada en (0,0) y suele estar rotada, por lo que a las coordenadas (x, y) en ecuación (4) de la elipse, se les debe rotar un ángulo θ de acuerdo con la ecuación (5). Así, el conjunto de pares coordenados (x', y') es el que se aproxima a las coordenadas 2D que forma la trayectoria del dedo corazón durante la ejecución del gesto 3.

$$\left(\frac{x^2 - c_x}{a}\right) + \left(\frac{y^2 - c_y}{b}\right) = 1 \quad (4)$$

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta \\ y' &= -x \sin \theta + y \cos \theta \end{aligned} \quad (5)$$

La elipse queda definida por el centro de masas de la elipse (c_x, c_y) , la rotación (θ) y la longitud de sus ejes mayor y menor (a, b) . Para hallar estos valores se realiza una aproximación, tal como se explica a continuación.

Para encontrar el centro de la elipse o centro de masas de forma aproximada, se parte del centroide y se realiza una búsqueda exhaustiva dentro de un cuadrado de 11 píxeles centrado en dicho centroide.

Durante la búsqueda, se estima la distancia euclidiana entre el centro de masa que se está probando y los puntos de la trayectoria. A partir de dichas distancias, se determinan las coordenadas (D_x, D_y) del punto con mayor distancia, el cual se asigna como eje mayor de la elipse; y, con esta información, se encuentra la inclinación aproximada de acuerdo con la ecuación (6).

$$\theta = \tan(Dx / Dy) \quad (6)$$

Posteriormente, se define el número de puntos que se acoplan al modelo de la elipse con los valores estimados. Para ello se compara la ordenada de la trayectoria con la ordenada de la elipse encontrada (definida por el centro de masa de prueba y, la inclinación y longitud de ejes estimada) dándole un margen de error de 5 píxeles.

Una vez terminada la búsqueda, se conservan los valores que definen la elipse con mayor número de puntos asociados y, al igual que en los casos anteriores, el número de puntos que se ajustan al modelo de la elipse define el porcentaje de éxito en la ejecución del gesto 3.

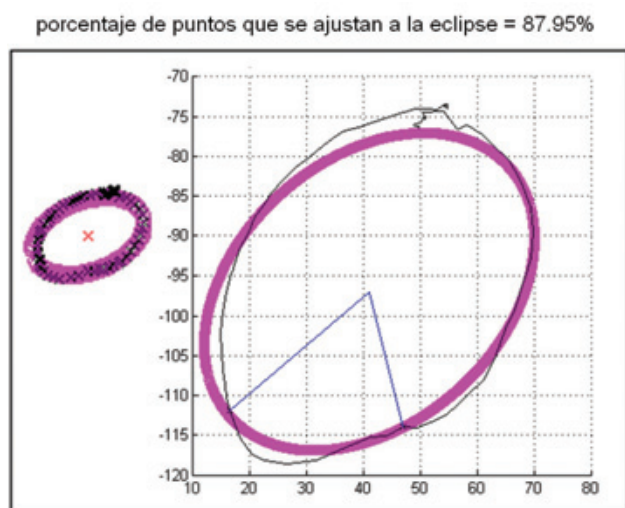


Fig. 4: Ejemplo de la métrica usada para el gesto 3

La Figura 4 muestra, en negro, la trayectoria seguida por la marca sobre el dedo corazón durante la ejecución del gesto 3 por parte de un demostrador. Los círculos fucsias (o) corresponden a la mejor elipse que encontró el sistema para ajustarse a los pares coordenados. La figura en el recuadro corresponde a una ampliación de la trayectoria, mientras que el fondo corresponde a la representación en el plano de la imagen.

3.5. MÉTRICA PARA EL GESTO DE ACERCAR Y ALEJAR LA MANO

El gesto 4 no tiene una forma particular en la trayectoria del efector final que se preservara durante las diferentes ejecuciones; de hecho, parece variar mucho con cambios pequeños en la orientación del demostrador hacia la cámara. Sin embargo, se encontró que el cambio en el tiempo del tamaño de la mano se puede aproximar a una *Campana de Gauss*.

Para estimar esta campana se empleó la función *fit* de Matlab. Se considera que el número de puntos que se ajustan al modelo son aquellos dentro de un rango de tolerancia del 15% de la diferencia entre el mayor y menor tamaño de la mano.

La Figura 5 presenta un ejemplo de este indicador durante la ejecución del gesto 4, en el eje, y corresponde al tamaño en píxeles de las mano y el eje x a la trama de video correspondiente. La línea azul corresponde al cambio en el tamaño de la mano original, los círculos negros (o) corresponden a la mejor Campana de Gauss encontrada por el sistema que se adapta a los datos, y las líneas rojas permiten visualizar el rango superior e inferior que acepta el sistema para considerar un punto como ajustado al modelo.

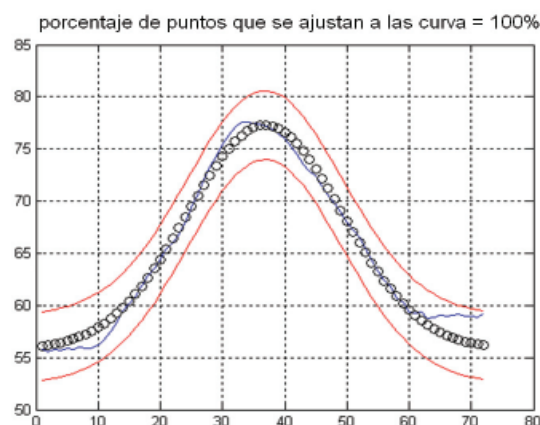


Fig. 5: Ejemplo de la métrica usada para el gesto 4

3.6. CRITERIOS GENERALES

Debe hacerse énfasis en que las formas son aproximadas y, rara vez se ajustan en un 100%, incluso cuando se usan los datos reales obtenidos durante las ejecuciones realizadas por el demostrador; sin embargo, se consideran una buena aproximación para identificar el éxito en la imitación, que no depende de la distancia a la cámara del ejecutor (invarianza a la escala), ni de la posición inicial, condiciones antropomórficas del demostrador, ni de la velocidad en la ejecución, entre otras condiciones claves, algunas indispensables, a la hora de generar una métrica.

4. RESULTADOS

La Tabla 1 resume los resultados de aplicar la métrica correspondiente al conjunto de 23 vídeos de prueba de cada gesto, seleccionados al azar.

Con la métrica usada para evaluar el gesto 1 (saludar), el menor porcentaje obtenido fue de 62 y el mayor de 100, mientras que el promedio fue de 86.03%. Un porcentaje bastante bueno si se tiene en cuenta que el margen de error permitido es bastante exigente (0.5 píxeles).

Con la métrica usada para evaluar el gesto 2 (abanicar), el menor porcentaje obtenido fue de 58.11, y el mayor de 100, mientras que el promedio fue de 79.47%; de acuerdo con esto se puede afirmar que, en general, se realizaron unas buenas ejecuciones del gesto 2. Debe tenerse en cuenta que, al igual que en caso anterior, el margen de error permitido fue bastante exigente (0.5 píxeles).

Vídeo	Métrica Gesto 1	Métrica Gesto 2	Métrica Gesto 3	Métrica Gesto 4
ejecución 1	100,00	78,87	61,45	79,16
ejecución 2	85,71	62,12	71,43	64,17
ejecución 3	92,11	58,11	87,76	56,30
ejecución 4	97,37	68,63	58,75	89,55
ejecución 5	88,57	90,16	70,27	82,00
ejecución 6	83,33	87,50	80,35	62,00
ejecución 7	78,57	75,31	76,74	96,07
ejecución 8	84,21	84,62	64,71	71,42
ejecución 9	81,69	91,93	81,63	86,95
ejecución 10	66,07	92,59	98,04	92,50
ejecución 11	71,70	63,51	97,57	47,56
ejecución 12	83,33	69,12	53,17	100,00
ejecución 13	86,84	61,40	71,01	93,87
ejecución 14	89,47	71,43	91,22	97,56
ejecución 15	80,00	69,09	79,45	60,34
ejecución 16	62,00	81,82	96,22	61,53
ejecución 17	90,00	100,00	59,09	60,71
ejecución 18	92,31	90,32	56,60	89,83
ejecución 19	91,89	98,15	70,42	47,36
ejecución 20	97,44	75,68	41,18	68,18
ejecución 21	97,37	85,37	51,25	100,00
ejecución 22	89,74	80,43	57,38	79,62
ejecución 23	88,89	91,63	75,00	89,25
PROMEDIO	86,03	79,47	71,77	77,21

Tabla 1: Resultados obtenidos por las cuatro métricas para 23 ejecuciones diferentes

Con la métrica usada para evaluar el gesto 3 (rotar la mano), el menor porcentaje obtenido fue de 41,18, y el mayor de 98,04, mientras que el promedio fue de 71,77%. El margen de error permitido en este caso fue de 5 píxeles.

La Figura 6 presenta un ejemplo que corresponde a uno de los casos por debajo del 70% establecido como límite; como puede observarse, dicho porcentaje fue provocado porque la trayectoria fue abierta. Sin embargo, el algoritmo encuentra una buena aproximación.

Con la métrica usada para evaluar el gesto 4 (acercar y alejar la mano), el menor porcentaje obtenido fue de 47,36 y el mayor de 100, mientras que el promedio fue de 77,21%. El margen de error permitido en este caso es proporcional a la diferencia entre el mayor y menor tamaño de la mano, es decir, al cambio en profundidad; la variación permitida fue del 15%.

Los resultados obtenidos con las métricas propuestas permiten concluir que las mejores ejecuciones ocurren en su orden: en el gesto 1 (86,03%), 2 (79,47%), 4 (77,21%) y 3 (71,77%).

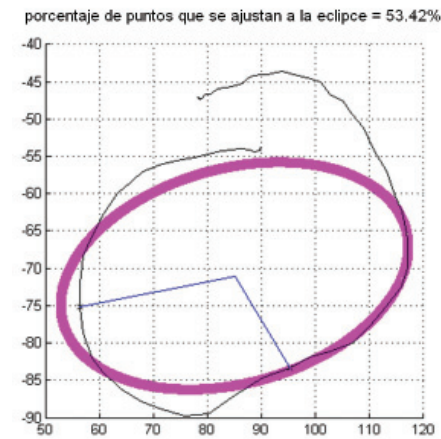


Fig. 6: Segundo ejemplo en detalle de la métrica usada para el gesto 3

4.1. MÉTRICAS VS ENCUESTAS

Con el objeto de contrastar los resultados de las métricas, se realizó una encuesta de dos preguntas a un grupo de ocho personas. El grupo de encuestados se conformó por 4 personas afines a las ciencias de la computación y 4 personas completamente ajenas a ellas.

Al principio se le pidió a los encuestados identificar el gesto ejecutado por el imitador, los resultados se resumen en la Tabla 2, donde los datos de la diagonal principal indican el porcentaje de veces que cada gesto fue correctamente identificado y, por lo tanto, lo ideal es que aparezca en la diagonal un valor de 100%; y cero fuera de ella. Los valores por fuera de la diagonal se interpretan como que el gesto identificado por la columna fue erróneamente identificado como el gesto dado por la fila.

Se observa que el Gesto mejor identificado fue el 3, seguido por el 2 y el 1. Por otro lado, el gesto peor identificado fue el 4, y la mayoría de las veces fue asociado a la opción **Otro**. El porcentaje de reconocimiento correcto varía entre el 39,67% y el 83,15%, con un promedio de reconocimiento del 61.96%.

	Gesto 1	Gesto 2	Gesto 3	Gesto 4	Otro
Gesto 1	61,96	17,93	1,63	14,13	4,35
Gesto 2	18,48	63,04	1,63	11,41	5,43
Gesto 3	7,61	1,63	83,15	3,26	4,35
Gesto 4	8,70	4,35	7,07	39,67	40,22

Tabla 2: Matriz de confusión promedio

Posteriormente, se le presentaron al encuestado las ejecuciones simultáneas del demostrador y del robot, seguidas de la pregunta: ¿Considera que el robot imitó el gesto realizado por el demostrador? Las respuestas posibles

eran: SI, NO. Los resultados a la segunda pregunta se presentan a través de la Figura 7.

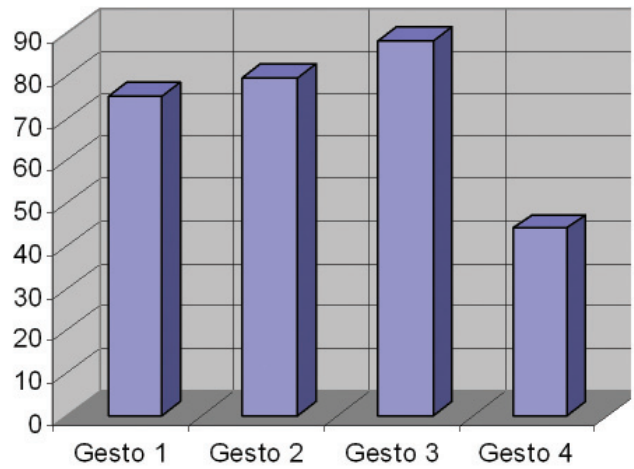


Fig. 7: Resultados a la Segunda Pregunta

El gesto mejor identificado fue el 3 (88,04%), seguido del gesto 2 (79,35%), gesto 1 (75%) y el peor reconocido fue el gesto 4 (44,02%). Los encuestados consideraron que hubo una buena imitación en el caso de los gestos 1 (saludar), 2 (abanicar) y 3 (rotar) con porcentajes superiores al 75%. Los resultados a esta pregunta son interesantes, pues permiten que el observador contraste las ejecuciones del demostrador con las del imitador, y, en muchos casos, este ejercicio hizo que cambiaran su opinión respecto a la buena ejecución del gesto evaluada en la primera pregunta. Esta diferencia se debe a que es posible no identificar correctamente el gesto y posteriormente, sin embargo, considerar que sí hubo una buena imitación por parte del robot. Este juicio, que parecía ser incongruente, se debió a la poca familiaridad de los encuestados con la forma en que se visualizó la ejecución de la imitación por el robot.

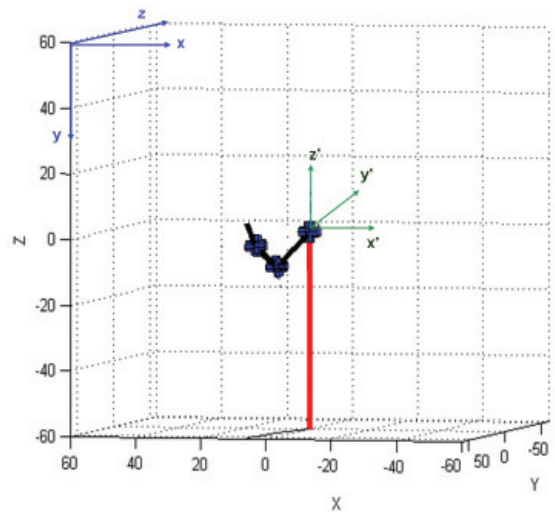


Fig. 8: Ejemplo de Visualización de la Ejecución por parte del Robot en un Instante de Tiempo

La Figura 8 presenta un ejemplo de visualización en un instante de tiempo del brazo robótico utilizando Matlab.

Esta afirmación se pudo corroborar a través de una segunda encuesta realizada a los mismos participantes, esta vez, con sólo 20 vídeos escogidos al azar, que incluían las imitaciones más deficientes. En este caso, se consideró la primera encuesta como un primer acercamiento a la visualización, y se pudo verificar que en la mayoría de los casos (6 de 8) hubo una notable mejoría en la correcta percepción del gesto interpretado, esto es, de un promedio de 61,96% a 68,75%. Adicionalmente, se notó que los encuestados que son afines a las ciencias de la computación percibieron con mayor facilidad los gestos interpretados por el robot y el 50% de este grupo mejoraron notablemente su desempeño en la segunda encuesta. En el grupo de personas no afines se encontraron los peores resultados de evaluación.

	Porcentajes de identificación	
	Métricas	Primera Encuesta
Gesto 1	86,03%	75,00 %
Gesto 2	79,47%	79,35 %
Gesto 3	71,77%	88,04 %
Gesto 4	77.21%	44,02 %

Tabla 3: Resumen de porcentajes de identificación

En la Tabla 3 se presenta el resumen de los resultados en la identificación, obtenidos a través de las métricas y el resultado de la primera encuesta. Se observa que, mientras que a través de las métricas la mejor imitación se presenta para el gesto 1, en la encuesta ocupa la tercera posición.

El gesto 2 ocupa el segundo puesto en los dos casos de análisis.

Aunque el gesto 3, a la luz del indicador es con el que peor resultados se obtiene, los encuestados consideraron que el gesto 3 tiene las mejores ejecuciones durante la imitación por parte del robot. Esto probablemente ocurre porque - para los encuestados - una buena imitación del gesto 3 incluye un rango de trayectorias más variado que el de una elipse rigurosa.

Lo contrario ocurre con el gesto 4 que, basándose en el indicador 4, parece tener un buen desempeño y, sin embargo, subjetivamente fue calificado en forma bastante deficiente (44,02%).

Se observa por tanto una mejor correspondencia entre los resultados de las métricas para los gestos 1, 2 y 3 con los resultados de la encuesta.

En general, la apreciación de los observadores se ve afectada por la poca percepción que se tiene a través de la visualización utilizada de los cambios en profundidad. Esta característica es especialmente relevante en la ejecución del gesto 4 - acercar y alejar la mano - y, probablemente, explican los resultados bajos dados por los evaluadores.

5. CONCLUSIONES Y TRABAJO FUTURO

Es posible asociar formas geométricas a la trayectoria seguida por un efector final con el ánimo de cuantificar la calidad de la imitación realizada por el robot. Esta propuesta resulta bastante simple y permite que el imitador ejecute su propia interpretación de un gesto observado. La propuesta no requiere que el demostrador y el imitador se encuentren en la misma escena, ni que posean las mismas características antropomórficas, e, incluso, pueden estar a diferentes distancias de la cámara y/o orientaciones relativas.

Se mejoró la evaluación sobre el resultado de la imitación, al calificar tanto cualitativa como cuantitativamente las ejecuciones del robot. En esta última evaluación se presentó un conjunto de métricas que tienen en cuenta la meta y tratan de despreciar, en menor o mayor grado, los detalles irrelevantes de la ejecución; además, son independientes de la trama de referencia y de la escala.

La programación de robots por demostración elimina la parametrización de trayectorias a realizar, auto-ajustándose a sus propias características morfológicas. Algunos campos posibles de aplicación son la robótica de servicios, de rehabilitación humana y en la construcción de interfaces hombre-máquina.

Como trabajo futuro se propone realizar un diseño experimental que considera la mayor cantidad de variables, estudiar la generación de formas geométricas en un espacio tridimensional. Así mismo, mejorar la visualización presentada a los encuestados de tal forma que los observadores puedan contar con una mejor percepción de las ejecuciones realizadas por el imitador.

6. AGRADECIMIENTOS

Agradecemos al Programa de apoyo a doctorados de Colciencias, a la Universidad del Valle y al Instituto Técnico Superior (IST) de Portugal, por el soporte a este trabajo. Un reconocimiento especial al profesor José Santos-Víctor del IST – Portugal por su orientación, consejo y apoyo a este proyecto.

7. BIBLIOGRAFÍA

- [1]Schaal S. 1999. 'Is imitation learning the route to humanoid robots?', Trends in Cognitive Sciences, vol. 3, no. 6, pp. 233-242.
- [2]Alissandrakis A, Nehaniv CL, Dautenhahn K & Saunders J. 2006. 'Evaluation of Robot Imitation Attempts: Comparison of the System's and the Human's Perspectives', 1st Annual Conference on Human-Robot Interaction (HRI), Salt Lake City, Utha, USA, pp. 134-141.
- [3]Billard A, Epars Y, Calinon S, Schaal S, & Cheng G. 2004, 'Discovering optimal imitation strategies'. Robotics and Autonomous Systems, Special Issue: Robot Learning from Demonstration, vol. 47, pp. 69-77.
- [4]Steinfeld A, Fong T, Kaber D, Lewis M, Scholtz J, Schultz A & Goodrich M. 2006. 'Common Metrics for Human-Robot Interaction', 1st Annual Conference on Human-Robot Interaction (HRI), Salt Lake City, Utha, USA, pp. 33-40.
- [5]Argall BD, Chernova S, Veloso M, & Browning B. 'A Survey of Robot Learning from Demonstration', Elsevier, Robotics and Autonomous Systems, 2005, vol 57, no. 5, pp. 469-483.
- [6]Mataric MJ & Pomplun M. 1998. 'Fixation behavior in observation and imitation of human movement', Cognitive Brain Research, vol. 7, pp. 191-202.
- [7]Thordike. Animal Intelligence, Macmillan, Ed. New York: Cited in (GALEF, 1988). Galef 1988, Imitation in animals: History, definition and interpretation of data from the psychological laboratory. Social learning: Psychological and Biological Perspectives, T. R. Z. B. G. G. (Eds.), Ed.: Hillsdale, NJ: Lawrence Erlbaum, pp. 3-28.
- [8]Nope S, Loaiza H & Caicedo, E. 'Modelo Bio-inspirado para el reconocimiento de gestos usando primitivas de movimiento en visión', Revista Iberoamerica de Automática e Informática Industrial (RIAI), 2008, vol. 5, no. 4, pp. 69-76.
- [9]Nope S, Loaiza H & Caicedo E. 'Reconstrucción 3D-2D de Gestos usando Información de Vídeo Monocular Aplicada a un Brazo Robótico', Revista Facultad de Ingeniería, Universidad de Antioquia, 2010, vol. 53.
- [10]Masters, T. 1995, Advanced Algorithms for Neural Networks a C++ Sourcebook.
- [11]Kuniyoshi Y, Yorozu Y, Inaba M & Inoue, H. 2003. 'From visuo-motor self learning to early imitation - a neural architecture for humanoid learning', International Conference on Robotics & Automation, Taipei, Taiwan. vol. 3, no. 3, pp. 3132-3139