

Jornadas de Automática

Control servo-visual híbrido para el posicionamiento de un manipulador robótico

Velasco, Edison^a, Neme, Juan Pablo^b, Gil, Pablo^{a,b,*}

^aInstituto Universitario de Investigación Informática, Universidad de Alicante, Crta. de San Vicente del Raspeig, s/n, 03690, San Vicente del Raspeig, España.

^bDepto. Física, Ing. de Sistemas y Teoría de la Señal, Universidad de Alicante, Crta. de San Vicente del Raspeig, s/n, 03690, San Vicente del Raspeig, España.

To cite this article: Velasco, E., Neme, J. P., Gil, P. 2026. Hybrid servo-visual control for the positioning of a Robotic Manipulator. Jornadas de Automática, 47. <https://doi.org/10.17979/ja-cea.2026.47.13752>

Resumen

Este trabajo presenta una arquitectura híbrida de control servo-visual para el posicionamiento de un manipulador robótico. La propuesta combina un controlador basado en posición (PBVS) y un controlador basado en imagen (IBVS) con una etapa final de aproximación por contacto. El controlador PBVS permite realizar un desplazamiento inicial hacia la pose objetivo utilizando información geométrica del sistema cámara-robot. Posteriormente, el controlador IBVS refina el posicionamiento mediante la minimización del error visual en el plano de la imagen, mejorando la precisión local frente a incertidumbres de calibración o pequeñas desviaciones en la pose. Finalmente, el robot realiza un movimiento controlado en dirección al punto deseado, hasta detectar contacto con el entorno mediante una condición de fuerza. La arquitectura se valida experimentalmente mediante tres escenarios: control PBVS, control IBVS y control híbrido PBVS-IBVS con aproximación por contacto. Los resultados permiten analizar la convergencia visual, la precisión de posicionamiento y la capacidad del sistema para completar una tarea de contacto de forma estructurada.

Palabras clave: Sistemas robóticos autónomos, Guiado y control, Manipuladores robóticos, Control de movimiento, Tecnología Robótica.

Hybrid servo-visual control for the positioning of a Robotic Manipulator

Abstract

This paper presents a hybrid servo-visual control architecture for positioning a robotic manipulator. The proposed approach combines a position-based controller (PBVS) and an image-based controller (IBVS) followed by a final stage of a contact-based approach. The PBVS controller enables an initial movement toward the target pose using geometric information from the camera-robot system. Subsequently, the IBVS controller refines the positioning by minimizing the visual error in the image plane to improve local accuracy despite calibration uncertainties or small pose deviations. Finally, the robot performs a controlled movement toward the desired point until it detects contact with the environment via a force condition. The architecture is experimentally validated using three scenarios: PBVS control, IBVS control, and hybrid PBVS-IBVS control with contact-based approach. The results enable analysis of visual convergence and positioning accuracy, as well as the system's ability to perform contact tasks in a structured manner.

Keywords: Autonomous robotic systems, Guidance and control, Robots manipulators, Motion control, Robotics Technology.

1. Introducción

En tareas de posicionamiento robótico guiadas por visión, los controladores servo-visuales permiten cerrar el lazo de control utilizando información proveniente de una cámara.

Entre las estrategias más usadas se encuentran el control basado en posición, PBVS, y el control basado en imagen, IBVS (Chaumette and Hutchinson, 2006). El primero permite definir el movimiento del robot a partir del error de pose entre el efector final y una referencia deseada. El segundo minimiza

*Autor para correspondencia: pablo.gil@ua.es
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

directamente el error entre características visuales observadas y deseadas en el plano de la imagen.

Las características extraídas de la imagen más habituales suelen ser: coordenadas de puntos de interés, tales como esquinas o centroides (Espiau et al., 1992) o marcadores fiduciales como ArUco o QR; momentos de imagen (Chaumette, 2004); líneas y segmentos; contornos o siluetas (Tahri and Chaumette, 2005); así como flujo óptico (Collewet and Marchand, 2011) o representaciones aprendidas por redes neuronales (Bateux et al., 2018). En este trabajo se ha decidido optar por el uso de esquinas y marcadores ArUco ya que ofrecen simplicidad computacional y buena convergencia local.

Aunque ambos enfoques de control han sido ampliamente estudiados, cada uno presenta limitaciones prácticas. PBVS depende de una correcta estimación geométrica y de una calibración precisa, mientras que IBVS puede presentar dificultades de convergencia cuando el error inicial es grande o cuando las características visuales se aproximan a los límites del campo visual. Estrategias híbridas como (Kermorgant and Chaumette, 2011) permiten superar estas limitaciones.

En este contexto, se propone una arquitectura de control servo-visual híbrida que combina ambas estrategias, PBVS e IBVS, para el posicionamiento de un brazo robótico sobre la superficie de un objeto. Esta arquitectura de control permite realizar una primera fase de aproximación global mediante PBVS, seguida de un posicionamiento local fino con IBVS y finalmente, un movimiento de aproximación final hasta que el efector robótico siente contacto con la superficie. La arquitectura se ha implementado en un brazo robótico Kinova Gen3 que hace uso de dos cámaras RGBD Intel RealSense y un sensor táctil PapillArray instalado en el efector final. Una de las cámaras se emplea en configuración *eye-to-hand* para la etapa PBVS, mientras que la otra cámara se monta en configuración *eye-in-hand* para la fase IBVS. El sensor táctil se utiliza en la fase final para detectar el contacto entre el efector robótico y el objeto.

El artículo se organiza de la siguiente manera. En la Sección 2 se presentan los fundamentos de los controladores servo-visual basados en imagen y en posición. En la Sección 3 se describe la arquitectura híbrida propuesta. La Sección 4 detalla la plataforma robótica, la configuración sensorial y la definición de las características visuales empleadas. En la Sección 5 se presentan los experimentos realizados y los resultados obtenidos. Finalmente, en la Sección 6 se exponen las conclusiones y los trabajos futuros.

2. Fundamentos de control servo-visual

2.1. Modelo de cámara y matriz de interacción

En este trabajo se considera el modelo de cámara pinhole (Sturm, 2021) para describir la proyección de las características visuales sobre el plano de imagen. Dado $\xi_c = [x_c, y_c, z_c]^T \in \mathbb{R}^3$ un punto característico expresado en el sistema de referencia de la cámara C en [mm], su proyección sobre el plano de imagen I se define como

$$\xi_I = \begin{bmatrix} u_I \\ v_I \end{bmatrix} = \frac{f}{z_c} \begin{bmatrix} x_c \\ y_c \end{bmatrix}, \quad (1)$$

donde $\xi_I = [u_I, v_I]^T \in \mathbb{R}^2$ representa la característica visual proyectada en la imagen medida en [pixel] y f es la distancia focal de la cámara. A partir de esta proyección, y de los parámetros intrínsecos de la cámara, la velocidad en imagen $\dot{\xi}$ de una característica visual ξ se puede relacionar con el movimiento de la cámara representado por su velocidad lineal y angular $\mathbf{v}_c = [v_x, v_y, v_z, \omega_x, \omega_y, \omega_z]^T$, como

$$\dot{\xi} = \mathbf{L}_s \mathbf{v}_c \quad (2)$$

donde $\mathbf{L}_s$ es la matriz de interacción o también conocida como Jacobiana de imagen.

2.2. Control IBVS

El control IBVS utiliza directamente las características observadas en el plano imagen en un instante de tiempo t para calcular la acción de control requerida para que éstas se desplacen en el mismo plano imagen hasta alcanzar una posición deseada en un tiempo finito $t + n$.

Por lo tanto, sea $\mathbf{s}$ el vector de características visuales actuales y $\mathbf{s}^*$ el vector de características deseadas, el error en la imagen es definido como $\mathbf{e}_I = \mathbf{s} - \mathbf{s}^*$. Si además se considera la relación de movimiento entre imagen y cámara dada por (2), entonces se puede imponer una dinámica decreciente sobre el error en la imagen, de forma que $\dot{\mathbf{e}}_I = -\lambda \mathbf{e}_I$, donde $\lambda > 0$ es una ganancia de control. De esta manera, la velocidad de la cámara se obtiene como

$$\mathbf{v}_c = -\lambda \mathbf{L}_s^+ \mathbf{e}_I, \quad (3)$$

donde $\mathbf{L}_s^+$ representa la pseudo-inversa de la matriz de interacción.

Esta ley de control permite calcular el movimiento necesario para reducir el error en imagen entre las características observadas en cada instante de tiempo y las deseadas. Notar que, para una configuración *eye-in-hand* en la que la cámara se monta en el efector, las velocidades calculadas en el sistema de referencia de la cámara se deben transformar al sistema de referencia utilizado por el controlador del brazo robótico.

El IBVS puede usarse para tareas de aproximación del efector final de un brazo robótico, ya que el error se corrige directamente en el plano de imagen. Sin embargo, su comportamiento puede verse afectado cuando: el error inicial es elevado, las características visuales se encuentran cerca de los límites del campo de visión, el objeto de referencia queda fuera del campo visual de la cámara, las características son colineales o si existen variaciones importantes de distancia z_c .

2.3. Control PBVS

El control PBVS define el error de control en el espacio cartesiano. A diferencia del IBVS, donde el error se calcula directamente en el plano de la imagen, en PBVS se estima la pose relativa entre el robot y una referencia deseada. Esta referencia puede estar asociada a la cámara, al efector final o al objeto observado, dependiendo de la configuración del sistema.

Sea $\mathbf{T} \in \text{SE}(3)$ la pose actual del efector final y $\mathbf{T}^* \in \text{SE}(3)$ la pose deseada. El error de pose se puede representar mediante una componente traslacional $\mathbf{t}_e$ y una componente rotacional $\boldsymbol{\theta}_e$, agrupadas en el vector $\mathbf{e}_P = [\mathbf{t}_e, \boldsymbol{\theta}_e]^T$.

A partir de este error, se define la siguiente ley de control proporcional

$$\mathbf{v}_r = -\mathbf{K}_p \mathbf{e}_p, \quad (4)$$

donde $\mathbf{K}_p$ es una matriz de ganancias positivas y $\mathbf{v}_r$ son las velocidades $\mathbf{v}_r = [v_x, v_y, v_z, \omega_x, \omega_y, \omega_z]^T$ en el sistema de referencia del robot. La componente traslacional permite reducir el error de posición, y la componente rotacional corrige la orientación del efector final respecto a la referencia deseada.

PBVS es recomendable cuando se requiere que un robot lleve a cabo desplazamientos grandes, ya que se emplean referencias de pose directamente en el espacio cartesiano. No obstante, su precisión depende de la estimación de la pose, de la calibración entre los sistemas de referencia involucrados y de la correcta transformación entre la cámara, el efector y la base del robot.

3. Sistema de posicionamiento visual y contacto

3.1. Control servo-visual híbrido

Los enfoques IBVS y PBVS presentan características complementarias. El control PBVS permite generar movimientos directos hacia una pose objetivo en el espacio cartesiano, por lo que es adecuado para realizar una aproximación inicial cuando el efector final de un brazo robótico se encuentra lejos de la referencia objetivo con la que tiene que contactar. Por otro lado, el control IBVS permite refinar el posicionamiento utilizando directamente el error observado en la imagen, esto resulta útil cuando el efector final ya se encuentra cerca de la referencia deseada.

En este trabajo se considera una estrategia híbrida en la que ambos controladores se emplean de forma secuencial, conmutando entre ambos. En primer lugar, el sistema utiliza PBVS para reducir el error de pose entre el efector final y la referencia objetivo. Cuando el error traslacional y rotacional se encuentra por debajo de umbrales definidos, el controlador cambia al modo IBVS. En esta segunda fase, se requiere minimizar el error en imagen entre las características observables y las características deseadas extraídas del objeto en la escena usado como referencia. La transición para pasar de PBVS a IBVS se ha definido mediante las siguientes dos condiciones:

$$\|\mathbf{t}_e\| < \epsilon_t, \quad \|\boldsymbol{\theta}_e\| < \epsilon_\theta, \quad (5)$$

donde ϵ_t y ϵ_θ son los umbrales de convergencia traslacional y rotacional, fijados en nuestros experimentos en 0,01 metros y 0,05 radianes, respectivamente. Una vez se cumplen estas dos condiciones, se desactiva el PBVS y activa el IBVS. La transición se regula mediante el error cartesiano de posición y orientación. Así, el IBVS se inicia cuando el efector se encuentra próximo a su referencia visual, reduciendo el riesgo de activarlo si se producen errores elevados o si la referencia se encuentra fuera del campo visual de la cámara *eye-in-hand*. Después, el controlador IBVS finalizará su tarea de guiado cuando el error en imagen cumpla, la siguiente condición:

$$\|\mathbf{e}_I\| < \epsilon_I, \quad (6)$$

donde ϵ_I es un umbral de las características de imagen, el cual está fijado a 5 píxeles en los experimentos. Además, esta condición debe mantenerse durante un intervalo de tiempo, evitando conmutaciones bruscas producidas por oscilaciones alrededor del umbral. Por tanto, la transición desde IBVS hacia la aproximación en z^+ solo se ejecuta cuando el error de las referencias visuales permanece por debajo de ϵ_I .

3.2. Aproximación por contacto

Una vez realizado el posicionamiento del brazo robot con referencias visuales, se requiere una última fase para conseguir que el brazo robot pueda interactuar por contacto con objetos en la escena. Esta fase consiste en desplazar el efector robótico en una dirección definida de su propio sistema de referencia, en este caso la dirección z^+ , hasta detectar la interacción física con un objeto del entorno mediante un sensor de fuerza situado en su extremo. La ejecución de esta fase de control permitirá llevar a cabo movimientos controlados de interacción con el objeto.

La detección de contacto se realiza a partir de una señal de fuerza medida F_z en la dirección del desplazamiento del efector. El controlador de aproximación por contacto finaliza el movimiento cuando se cumple que:

$$F_z \geq F_{th}. \quad (7)$$

donde F_{th} es el umbral definido para detectar contacto.

En este trabajo, esta fase no se plantea como un control de interacción continuo como se muestra en (Velasco-Sanchez et al., 2025), sino como una estrategia de aproximación y detección inicial de contacto. Por esta razón, no se implementa un controlador de admitancia o impedancia, ya que el objetivo de la arquitectura propuesta es validar la transición desde el posicionamiento visual hasta el primer contacto con la superficie.

4. Plataforma robótica para posicionamiento

4.1. Configuración de pruebas

El control servo-visual híbrido con fase final de aproximación por contacto descrito en la sección anterior, se ha implementado en ROS 2 Humble utilizando un brazo robótico Kinova Gen3 de 7 grados de libertad, dos cámaras RGBD Intel RealSense (D415 y D435) y un sensor táctil PapillArray (Khamis et al., 2019) instalado en el efector final. La configuración general del sistema se muestra en la Figura 1.

La cámara D415 se encuentra fija respecto al entorno de trabajo, siguiendo una configuración *eye-to-hand*. Esta cámara observa globalmente el área de trabajo y se utiliza durante la fase PBVS para estimar la referencia cartesiana del robot. La cámara D435 está montada sobre el efector final, siguiendo una configuración *eye-in-hand*. Esta cámara se utiliza durante la fase IBVS para refinar el posicionamiento a partir de características visuales observadas localmente desde el efector final.

Para la obtención de parámetros intrínsecos de la cámara *eye-in-hand*, que determinan la matriz de calibración, se ha usado el algoritmo de calibración de Brown–Conrady disponible en (Tirado-Garín and Civera, 2025). Y, para la obtención

de los parámetros extrínsecos de esa misma cámara se ha empleado el método de (Tsai and Lenz, 1989) implementado en el paquete software *easy_handeye2* (Free Software Foundation Inc., 2021). Una vez conocido el modelo de la cámara, así como conocidas las transformaciones entre la pose de la cámara anclada al efector y el propio efector final del robot, es posible determinar la posición y orientación de los objetos observados por la cámara respecto al sistema de referencia del robot. Del mismo modo, para la cámara *eye-to-hand*, se obtuvo la transformación entre la cámara fija y la base del robot. Todas estas transformaciones rígidas se integran en el árbol de transformaciones de ROS 2 y permiten expresar las mediciones visuales y los comandos de velocidad en los sistemas de referencia requeridos por cada fase de control.

El sensor táctil PapillArray se ha instalado con la configuración usada en (Velasco-Sanchez et al., 2025), y se utiliza únicamente durante la etapa final de aproximación por contacto. Una vez en las fases servo-visuales se cumplen (5) y (6), el robot avanza en la dirección z^+ del efector final hasta que se cumple (7).



Figura 1: Plataforma robótica experimental sobre la que se ha implementado el controlador servo-visual híbrido para posicionamiento sobre un objeto

4.2. Detección y definición de características visuales

En la experimentación se han utilizado marcadores ArUco (Garrido-Jurado et al., 2014) como referencias visuales para los controladores servo-visuales. Notar que el método propuesto no hace uso del canal de profundidad de las cámaras RGBD y, por lo tanto, la pose relativa del marcador ArUco se estima a partir de información monocular y de la calibración de las cámaras. Destacar, también, que el marcador se emplea de forma diferente en cada fase del controlador híbrido propuesto. Así, en PBVS se utiliza su pose tridimensional estimada respecto a la cámara externa fija, y luego se transforma al sistema de referencia de la base del robot para definir la referencia cartesiana del efector final. Mientras que en IBVS se utilizan marcas fiduciales directamente obtenidas desde las cuatro esquinas externas del ArUco, detectadas en la imagen desde la cámara montada en el efector final.

Durante la fase PBVS, la cámara *eye-to-hand* detecta el marcador ArUco y estima su pose mediante funciones implementadas en OpenCV y el paquete software (PAL Robotics S.L., 2020), utilizando los parámetros intrínsecos de la cámara obtenidos en el proceso de calibración y el tamaño conocido

del marcador ArUco. De esta forma, la distancia entre cámara y marcador se da mediante la componente traslacional de la pose estimada, sin requerir del canal de profundidad de la cámara RGBD. Después, la pose estimada en imagen del marcador ArUco se transforma al sistema de referencia de la base del robot, a partir de los parámetros extrínsecos previamente obtenidos. Esta información permite definir la referencia cartesiana hacia la cual debe aproximarse el efector final del robot mediante (4). Conviene notar que nuestro PBVS realiza un control basado en posición y en orientación desacoplado, es decir posición y orientación del efector se controlan por separado. Esto permite ajustar de manera independiente las ganancias de traslación y rotación del controlador, reduciendo movimientos angulares bruscos durante la aproximación global.

Durante la fase IBVS, la pose 3D completa del marcador no se utiliza como variable de error del controlador. En su lugar, se emplean las coordenadas en imagen de las cuatro esquinas del marcador detectado por la cámara *eye-in-hand*. Para construir la matriz de interacción se requiere una profundidad asociada a cada característica visual; en esta implementación, dicha profundidad se aproxima a partir del componente traslacional de la pose estimada del marcador respecto a la cámara. De esta forma, el vector de características visuales actuales se define como $\mathbf{v}_r = [u_1, v_1, u_2, v_2, u_3, v_3, u_4, v_4]^T$, donde (u_i, v_i) representa la posición de la esquina i del marcador ArUco en el plano de imagen. Entonces, el controlador IBVS compara estas características actuales con un conjunto de características deseadas predefinidas $\mathbf{s}^*$, como se mostró en Sección 2.2. Estas características deseadas representan la configuración objetivo del marcador en la imagen de la cámara *eye-in-hand*. Por tanto, durante la fase IBVS, el objetivo del controlador es desplazar las esquinas observadas del marcador hasta hacerlas coincidir con sus posiciones deseadas en la imagen mediante la ley de control (3).

4.3. Método híbrido PBVS–IBVS–Contacto

El método implementado se organiza como una secuencia de tres fases: PBVS, IBVS y aproximación por contacto, tal como se muestra en la Figura 2. Cada fase utiliza una fuente de información diferente: la cámara *eye-to-hand* para el posicionamiento cartesiano en SE(3), la cámara *eye-in-hand* para el refinamiento visual y el sensor PapillArray para la detección de contacto.

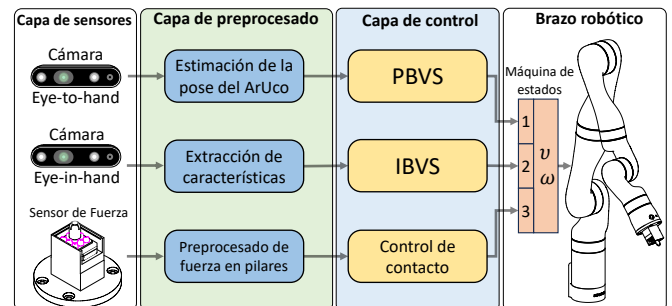


Figura 2: Arquitectura de control implementada para posicionamiento sobre la superficie de un objeto.

La arquitectura de control está organizada de la siguiente

manera. Inicialmente, la plataforma robótica ejecuta el controlador PBVS para aproximar el efector final a una pose objetivo calculada a partir de la pose del marcador ArUco observado por la cámara fija. En el instante donde el error cartesiano cumple los umbrales definidos como en (5), el sistema conmuta para ejecutar el controlador IBVS. En esta fase, se utilizan las cuatro esquinas del marcador detectadas por la cámara montada en el efector final para reducir el error visual respecto a una configuración deseada y la distancia que se tiene de cámara a marcador ArUco. Después, una vez que el error visual cumple (6) y se mantiene por debajo de este umbral, se activa la fase de aproximación por contacto. En esta última fase, el efector final avanza en la dirección z^+ de su propio sistema de referencia hasta que la fuerza medida por el sensor PapillArray supera el umbral de contacto y se mantiene cumpliendo (7). Finalmente, el robot detiene el movimiento y la tarea de aproximación se considera completada.

5. Experimentos

Los experimentos se realizaron con la configuración descrita en la Sección 4. El objetivo de esta evaluación es analizar el comportamiento de los controladores PBVS e IBVS, tanto de forma independiente como integrados dentro de una arquitectura híbrida PBVS-IBVS-Contacto.

En todos los experimentos, el marcador ArUco utilizado como referencia permanece estático; por tanto, el seguimiento de objetos móviles queda fuera del alcance de este trabajo.

5.1. Evaluación de los controladores PBVS e IBVS

Los controladores PBVS e IBVS se evaluaron en cuatro tipos de pruebas, en las que se modificó la configuración espacial del marcador ArUco utilizado como referencia visual. Cada prueba está compuesta por cinco experimentos de posicionamiento. En todos los casos, el robot parte de la misma configuración inicial, mientras que la posición final del ArUco varía de acuerdo con los siguientes casos:

Prueba 1: variación en el plano XY. El ArUco mantiene la misma orientación que la base del robot en los ejes X, Y y Z, pero se sitúa en posiciones distintas dentro del plano XY.

Prueba 2: variación en XY y altura Z. El ArUco mantiene la misma orientación que la base del robot en los ejes X e Y, pero se modifica su posición en el eje Z. Además, se consideran distintas posiciones en el plano XY al mismo tiempo.

Prueba 3: variación en XY y yaw. El ArUco se sitúa en distintas posiciones dentro del plano XY y se introduce una variación de orientación alrededor del eje Z, correspondiente al ángulo yaw.

Prueba 4: variación general en posición y orientación. El ArUco se sitúa en diferentes posiciones dentro del espacio XYZ y con distintas orientaciones respecto a la base del robot.

5.2. Resultados experimentales

Los resultados obtenidos para los controladores PBVS e IBVS se muestran en la Fig. 3. En esta evaluación se analiza principalmente la evolución temporal de las variables de error y de las velocidades generadas por cada controlador durante las pruebas realizadas.

En el caso del controlador PBVS, se observa que los errores cartesianos de posición y orientación disminuyen progresivamente durante la ejecución de los experimentos. De forma similar, las velocidades lineales y angulares generadas por el controlador tienden a valores cercanos a cero conforme el efector final se aproxima a la referencia deseada. Para todos los experimentos realizados con PBVS, las variables analizadas alcanzan valores próximos a cero en un tiempo inferior a 10 [s], lo que muestra la capacidad del controlador para realizar la aproximación global hacia la pose objetivo.

En el caso del controlador IBVS, el error visual asociado a las características del marcador ArUco también disminuye a lo largo del tiempo. Las coordenadas de las esquinas detectadas convergen hacia sus posiciones deseadas en la imagen, mientras que las velocidades calculadas por el controlador se reducen progresivamente hasta valores cercanos a cero. En las pruebas realizadas, el controlador IBVS alcanza una condición próxima a la convergencia alrededor de los 8 [s].

Los resultados muestran que, en todos los experimentos, ambos controladores tienen un comportamiento de convergen-

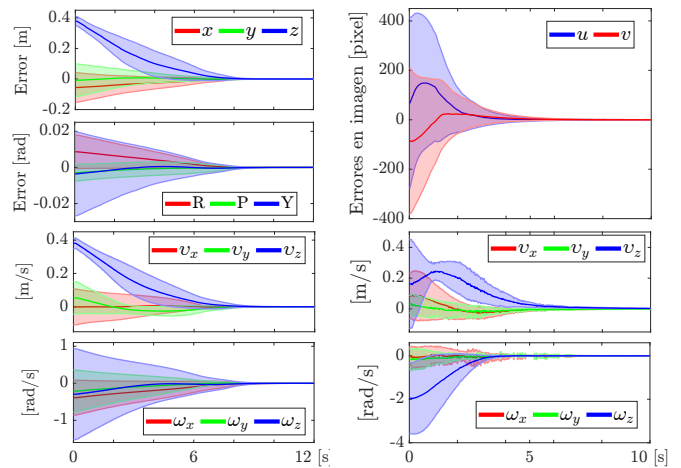


Figura 3: Evolución temporal de los errores y velocidades. (izquierda) Controlador PBVS. (derecha) Controlador IBVS.

5.3. Experimento con controlador híbrido PBVS-IBVS-Contacto

Además de evaluar los controladores PBVS e IBVS de forma independiente, se realizó un experimento para validar la arquitectura híbrida completa. En esta prueba, el sistema ejecuta una máquina de estados compuesta por las tres fases: aproximación global mediante PBVS, refinamiento visual mediante IBVS y aproximación final por contacto.

Durante la primera fase, el robot utiliza el controlador PBVS para aproximar el efector final hacia la referencia estimada a partir del marcador ArUco observado por la cámara *eye-to-hand*. Una vez que el error cartesiano alcanza los umbrales definidos en (5), que se consiguen aproximadamente entorno a los 10 [s], el sistema cambia al modo IBVS. En esta segunda fase, la cámara *eye-in-hand* detecta el marcador ArUco y el controlador reduce el error visual entre las coordenadas en imagen de las esquinas detectadas del marcador y sus valores deseados. Finalmente, cuando el error visual se mantiene

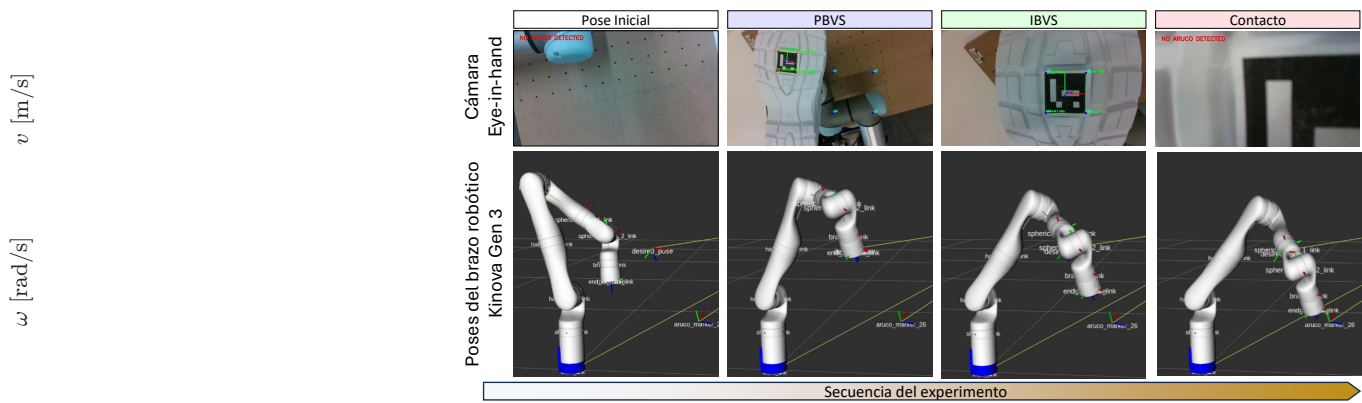


Figura 4: Experimentos con el controlador híbrido PBVS-IBVS-Contacto. (izquierda) Velocidades lineales y angulares generadas durante la ejecución. (derecha) Secuencia experimental observada desde la cámara *eye-in-hand* y visualización de la pose del robot Kinova Gen3 durante cada fase del controlador.

por debajo del umbral establecido en (6) durante un intervalo mínimo de tiempo, que se alcanza aproximadamente a los 8.2 [s], se activa la fase de aproximación por contacto, en la que el efector final avanza en la dirección z^+ de su propio sistema de referencia hasta que el sensor PapillArray proporciona una magnitud de fuerza media mayor a 1 [N], en cuyo instante se da por terminada la tarea de posicionamiento.

En la Fig. 4 se muestran las velocidades lineales y angulares generadas durante la ejecución de la tarea, además de la secuencia experimental. El PBVS presenta las velocidades de mayor magnitud, asociadas al movimiento de aproximación global. Posteriormente, durante la fase IBVS, las velocidades disminuyen progresivamente hasta valores cercanos a cero, lo que indica la convergencia del refinamiento visual. En la fase final de aproximación por contacto, las velocidades angulares permanecen prácticamente nulas y se mantiene únicamente el movimiento lineal de aproximación hasta la detección del contacto.

6. Conclusiones y trabajos futuros

En este trabajo se presentó una arquitectura de control servo-visual híbrida para el posicionamiento de un brazo robótico sobre superficies. Los resultados experimentales muestran que los controladores visuales evaluados presentan un comportamiento convergente en las condiciones de prueba consideradas, permitiendo reducir progresivamente los errores de posición y visuales antes de iniciar la aproximación por contacto. La integración secuencial de las etapas de control permitió completar la tarea de posicionamiento, desde la aproximación inicial hasta la detección del contacto con la superficie mediante el sensor táctil.

Como trabajo futuro, se plantea extender la arquitectura propuesta eliminando la dependencia de marcadores ArUco. Para ello, será necesario incorporar métodos de detección de objetos, segmentación visual o estimación de pose 6D que permitan definir automáticamente las referencias de la tarea y los sistemas de coordenadas asociados al objeto y a la superficie de contacto.

Agradecimientos

Este trabajo ha sido financiado por Interreg-VI Sudoe y el European Regional Development Funds a través del proyecto europeo REMAIN (Ref. S1/1.1/E0111).

Referencias

- Bateux, Q., Marchand, E., Leitner, J., Chaumette, F., Corke, P., 2018. Training deep neural networks for visual servoing. In: IEEE Int. Conf. on Robotics and Automation (ICRA). pp. 3307–3314.
DOI: 10.1109/ICRA.2018.8461068
- Chaumette, F., 2004. Image moments: a general and useful set of features for visual servoing. IEEE Trans. on Robotics 20 (4), 713–723.
DOI: 10.1109/TR0.2004.829463
- Chaumette, F., Hutchinson, S., 2006. Visual servo control. i. basic approaches. IEEE Robotics & Automation Magazine 13 (4), 82–90.
DOI: 10.1109/MRA.2006.250573
- Collewet, C., Marchand, E., 2011. Photometric visual servoing. IEEE Trans. on Robotics 27 (4), 828–834.
DOI: 10.1109/TR0.2011.2112593
- Espiau, B., Chaumette, F., Rives, P., 1992. A new approach to visual servoing in robotics. IEEE Trans. on Robotics and Automation 8 (3), 313–326.
DOI: 10.1109/70.143350
- Free Software Foundation Inc., 2021. easy_handeye2: automated, hardware-independent hand-eye calibration for ros2.
URL: https://github.com/marcoesposito1988/easy_handeye2/
- Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., Medina-Carnicer, R., 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognition 47 (6), 2280–2292.
- Kermorgant, O., Chaumette, F., 2011. Combining ibvs and pbvs to ensure the visibility constraint. In: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems. pp. 2849–2854.
DOI: 10.1109/IR0S.2011.6094589
- Khamis, H., Xia, B., Redmond, S., 2019. A novel optical 3d force and displacement sensor-towards instrumenting the papillarray tactile sensor. Sensors and Actuators A: Physical 291, 174–187.
DOI: 10.1016/j.sna.2019.03.051
- PAL Robotics S.L., 2020. Aruco ros.
URL: https://github.com/pal-robotics/aruco_ros
- Sturm, P., 2021. Pinhole camera model. In: Ikeuchi, K. (Ed.), Computer Vision. A Reference Guide. Springer Cham, pp. 983–986.
DOI: 10.1007/978-3-030-63416-2
- Tahri, O., Chaumette, F., 2005. Point-based and region-based image moments for visual servoing of planar objects. IEEE Trans. on Robotics 21 (6), 1116–1127.
DOI: 10.1109/TR0.2005.853500
- Tirado-Garín, J., Civera, J., 2025. Anycalib: On-manifold learning for model-agnostic single-view camera calibration. In: IEEE/CVF Int. Conf. on Computer Vision (ICCV). pp. 8044–8055.
DOI: 10.1109/ICCV51701.2025.00754
- Tsai, R., Lenz, R., 1989. A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. IEEE Trans. on Robotics and Automation 5 (3), 345–358.
DOI: 10.1109/70.34770
- Velasco-Sanchez, E., Castaño-Amoros, J., Gil, P., Torres, F., 2025. Touch-based effector control to track 3d surfaces. In: IEEE 30th Int. Conf. on Emerging Technologies and Factory Automation (ETFA). pp. 1–6.
DOI: 10.1109/ETFA65518.2025.11205606