

Instituto Tecnológico de Costa Rica
Área Académica de Ingeniería Mecatrónica



Implementación de un sistema para la asistencia de pacientes mediante el robot Pepper

para optar por el título de
Ingeniero en Mecatrónica
con el grado académico de
Licenciatura

Steven Antonio Jiménez Bustamante

Cartago, 2 de octubre de 2017

Declaro que el presente informe Proyecto de Graduación ha sido realizado enteramente por mi persona, utilizando y aplicando literatura referente al tema e introduciendo conocimientos propios.

En los casos en que he utilizado bibliografía he procedido a indicar las fuentes mediante las respectivas citas bibliográficas. En consecuencia, asumo la responsabilidad total por el trabajo de graduación realizado y por el contenido del correspondiente informe final.

Steven Antonio Jiménez Bustamante

Cartago, 2 de octubre de 2017

Céd: 1-1579-0679

This work is licensed under a [Creative Commons](#)
“[Attribution-NonCommercial-ShareAlike](#) 4.0 [International](#)”
license.



Tecnológico de Costa Rica
Área Académica de Ingeniería Mecatrónica
Proyecto de Graduación
Tribunal Evaluador

Proyecto de Graduación defendido ante el presente Tribunal Evaluador como requisito para optar por el título de Ingeniero en Mecatrónica con el grado académico de Licenciatura, del Tecnológico de Costa Rica.

Miembros del Tribunal

M.Sc. Ing. Gabriela Ortiz León
Profesora Lectora

M.Sc. Ing. Sergio Arriola Valverde
Profesor Lector

M.Sc. Ing. Yeiner Arias Esquivel
Profesor Asesor

Los miembros de este Tribunal dan fe de que el presente trabajo de graduación ha sido aprobado y cumple con las normas establecidas por Área Académica de Ingeniería Mecatrónica.

Cartago, 15 de enero de 2018

Tecnológico de Costa Rica
Área Académica de Ingeniería Mecatrónica
Proyecto de Graduación
Tribunal Evaluador
Acta de Evaluación

Proyecto de Graduación defendido ante el presente Tribunal Evaluador como requisito para optar por el título de Ingeniero en Mecatrónica con el grado académico de Licenciatura, del Tecnológico de Costa Rica.

Estudiante: Steven Antonio Jiménez Bustamante

Nombre del Proyecto: *Implementación de un sistema para la asistencia de pacientes mediante el robot Pepper*

Miembros del Tribunal

M.Sc. Ing. Gabriela Ortiz León
Profesora Lectora

M.Sc. Ing. Sergio Arriola Valverde
Profesor Lector

M.Sc. Ing. Yeiner Arias Esquivel
Profesor Asesor

Los miembros de este Tribunal dan fe de que el presente trabajo de graduación ha sido aprobado y cumple con las normas establecidas por el Área Académica de Ingeniería Mecatrónica.

Nota final del Proyecto de Graduación: _____

Cartago, 2 de octubre de 2017

Resumen

Resumen Español

Palabras clave: Visión por Computadora, ADAS, Aproximada Visión por computador, ROS, Robot Pepper, Faster R-CNN, Redes Neuronales Convolucionales, Deep Learning

Abstract

Abstrac English

Keywords: Computer Vision, ADAS, Aproximate

a mis queridos padres

Índice general

Índice de figuras	iii
Índice de tablas	v
1 Introducción	1
1.1 Contexto: Situación actual en Costa Rica	1
1.2 Deep Learning	5
1.3 Datasets	5
1.4 ROS	5
1.5 Deep Learning Frameworks	5
1.6 Metodología utilizada en la investigación	5
1.7 Objetivos de la investigación	6
1.8 Estructura del documento	6
2 Integracion de tareas	7
2.1 Reconocer los objetos especificados en el entorno que rodea al robot:	7
2.2 Reconocer la voz del paciente y traducirlo como una instrucción:	8
2.3 Integrar las tareas de desplazamiento y el movimiento de objetos por parte del robot junto con el reconocimiento de objetos y de voz:	8

Índice de figuras

1.1	Fotografía del Robot Pepper. [?]	2
1.2	Diagrama de Solucion	4

Índice de tablas

Capítulo 1

Introducción

link http://brohrer.github.io/how_convolutional_neural_networks_work.html

<http://cs231n.github.io/convolutional-networks/architectures>

<https://adeshpande3.github.io/A-Beginner>

<http://cs231n.github.io/>

<http://www.datosgratis.net/consejos-para-la-introduccion-de-una-tesis/>

1.1 Contexto: Situación actual en Costa Rica

Es bastante claro que, en todos los países, donde Costa Rica no es la excepción, existen personas que por razones de enfermedad, accidentes, vejez u otras situaciones, son declaradas con algún tipo de discapacidad y que de manera parcial o total pierden la movilidad en su cuerpo.

A manera de ejemplo y basados en el Vigésimo Segundo Informe Estado de la Nación en Desarrollo Humano Sostenible de Costa Rica [?], en el capítulo 5 referente al Fortalecimiento de la Democracia se expresa lo siguiente:

“De acuerdo con el censo de 2011, en Costa Rica viven 452.849 personas con algún tipo de discapacidad, lo que representa alrededor de un 10% de la población total.”

Esto hace pensar que una cantidad importante de personas en Costa Rica, y de hecho en todos los países del mundo, sufren de alguna discapacidad y basado en esto se pueden proponer soluciones a este problema que les garantice una mejor calidad de vida.

Una de las principales prioridades de los profesionales, en este caso, la ingeniería, es buscar el bienestar de los demás y lograr una vida digna para ellos, es por esto que el problema que afronta este proyecto es la dificultad de ejecución de actividades cotidianas

por parte de personas con algún tipo de discapacidad, y en este caso se enfoca en la tarea de alcanzar objetos que se encuentran en su entorno.

Aunque este trabajo puede ser ejecutado por personas capacitadas en esta área, esta tarea también puede ser ejecutada por robots humanoides, capaces de cumplir con labores varias, y reubicar a este personal en actividades que requieran habilidades más complejas para mejorar la atención a los pacientes. Por esta razón se propone implementar un sistema para la asistencia de pacientes mediante el robot Pepper.

La posibilidad de implementar un robot humanoide que pueda asistir al paciente facilita el cuidado que requieren estas personas. Debido a que el proyecto se hará utilizando el robot humanoide Pepper, no existen opciones de selección de hardware, sino más bien, utilizando todo el potencial de este robot utilizar todas las herramientas de software para ejecutar este proyecto de la mejor manera. Los principales sistemas que posee este robot son sensores como una unidad inercial, láseres, sensores infrarrojos, sensores ultrasónicos, codificadores rotativos magnéticos y sensores táctiles y de contacto; los sistemas de interacción contemplan altavoces, micrófonos, cámaras 2D, sensor 3D (dispositivo tipo Kinect), LEDs, además los sistemas de movimiento y traslación contemplan varios motores.

A continuación, se muestra en la Figura 1.1 una fotografía del robot Pepper donde se pueden observar algunos de los módulos mencionados anteriormente.



Figura 1.1: Fotografía del Robot Pepper. [?]

Utilizando todas estas herramientas que posee el robot, el problema se puede solucionar de la siguiente manera:

El robot debe estar pendiente de las instrucciones que le indique el paciente, para lo cual, se utilizarán los micrófonos para grabar el audio (mensaje) emitido por el paciente, posteriormente convertirlo a texto y por último interpretar este texto como una instrucción. Después de interpretar la instrucción, el robot debe buscar el objeto que le pidió

el paciente mediante el uso de redes neuronales convolucionales de región (en inglés r-cnn: region convolutional neural network). Si el robot no logra localizarlo, éste deberá notificarle al paciente que no encontró el objeto; por otro lado, si el robot logra localizar el objeto, éste deberá desplazarse hasta donde está el objeto, sujetar el objeto y llevarlo donde se encuentra el paciente. El robot debe entregarle el objeto al paciente en un lugar cercano, por ejemplo, una mesa de noche, con el objetivo de velar por la salud y seguridad del paciente. Para términos prácticos el entorno donde se llevará a cabo las pruebas y evaluación del proyecto será en el laboratorio de Robótica y Visión Tridimensional en la Universidad de Alicante, España. En la figura 1.2 se puede brindar un diagrama de flujo que facilita la comprensión de la propuesta para hacer frente al problema en cuestión.

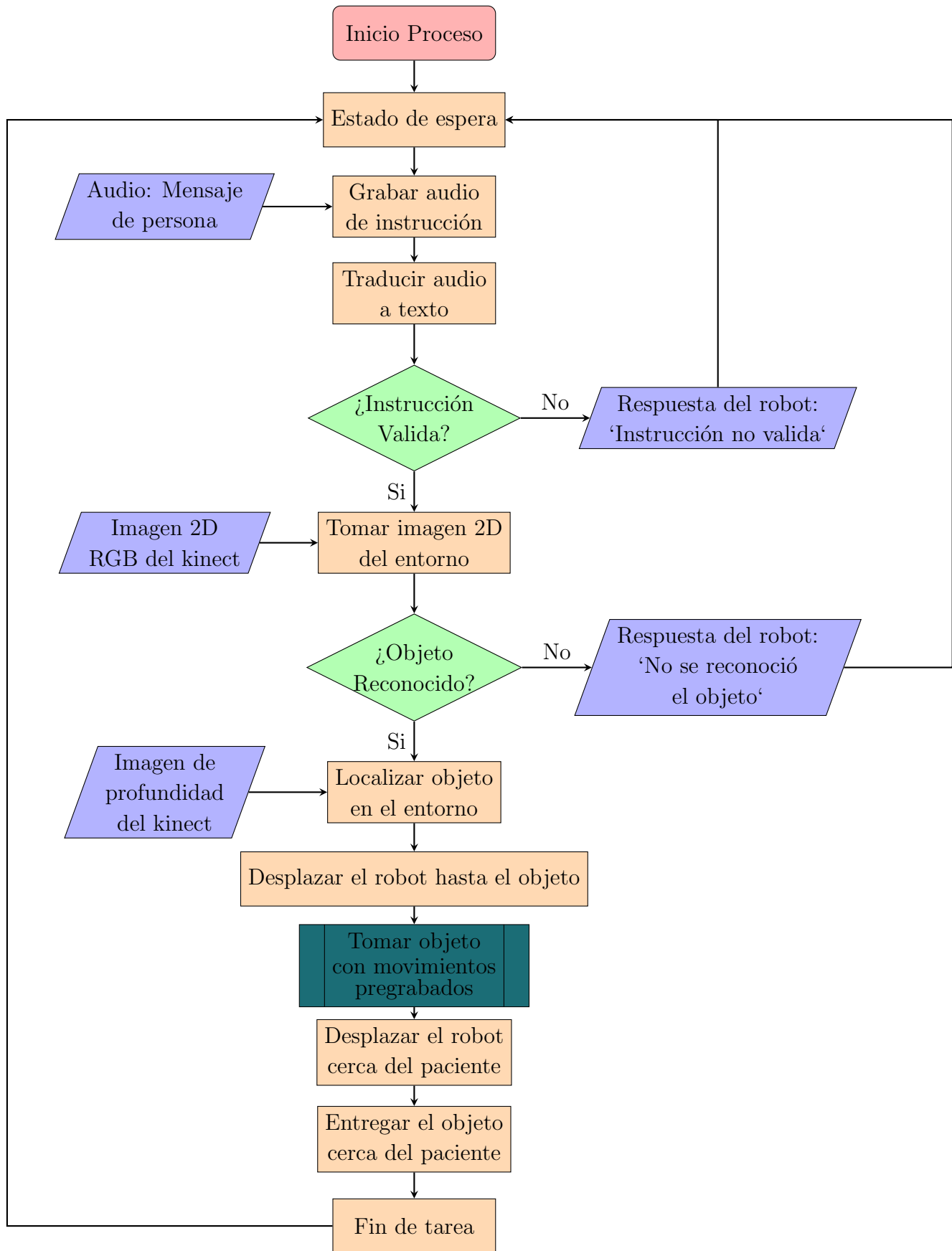


Figura 1.2: Diagrama de Solucion

1.2 Deep Learning

agregar un poco sobre CNN y RPN, luego en marco teorico se amplia

1.3 Datasets

1.4 ROS

1.5 Deep Learning Frameworks

1.6 Metodología utilizada en la investigación

La metodología utilizada en el proceso de investigación debe ser descrita en forma breve o resumida, de tal manera que el lector comprenda cómo se hizo la tesis y qué elementos fueron utilizados para demostrar o negar la hipótesis (diseño de la investigación, tipo de muestreo empleado, tamaño de la muestra, instrumentos empleados para la recopilación de información, etc.).

1.7 Objetivos de la investigación

1. Objetivo general:

Diseñar un sistema que asista al paciente en la tarea específica de alcanzar objetos según lo requiera la persona utilizando el robot Pepper

2. Objetivos específicos:

2.1. Implementar el reconocimiento de los objetos específicos en el entorno que rodea al robot.

2.1.1. Indicador: El robot debe reconocer tres objetos pequeños mínimo el 90% de los intentos en un rango de 1 a 5 metros de distancia.

2.2. Implementar el reconocimiento de la voz del paciente y traducirlo como una instrucción.

2.2.1. Indicador: El robot debe reconocer la siguientes palabras: traer, alcanzar y los tres nombres de los objetos que se deben reconocer, mínimo un 90% de los intentos cuando esté a 1 metro de distancia o menos del paciente, una vez que el mensaje haya sido traducido de audio a texto.

2.3. Integrar las tareas de desplazamiento y el movimiento de objetos por parte del robot junto con el reconocimiento de objetos y de voz.

2.3.1. Entregable: Esta integración debe cumplir con las métricas definidas anteriormente además de solucionar el problema de la siguiente forma: Cuando la instrucción haya sido interpretada, el robot debe buscar el objeto deseado, determinar su ubicación y desplazarse hasta este a una distancia de aproximadamente 35 centímetros, posteriormente deberá tomarlo y llevarlo hasta una posición cerca del paciente de igual manera a una distancia aproximada de 35 centímetros. Si el robot no logra ubicar el objeto, deberá indicarle al paciente que no logró identificarlo.

1.8 Estructura del documento

Respecto a la explicación breve del contenido de la investigación se debe narrar en qué consiste cada capítulo de la tesis de tal manera que el lector decida continuar leyendo toda la tesis o decida ignorarla por no ser un tema de su interés.

Capítulo 2

Integracion de tareas

Procedimiento para la ejecución del proyecto

Este proyecto reúne distintas disciplinas como visión artificial (computer vision), aprendizaje profundo (deep learning), integración de tareas mediante software tipo middleware (lógica de intercambio de información entre aplicaciones), nubes de puntos, entre otros. A continuación se explica las actividades que se pretenden aplicar para lograr cada etapa planteada en los objetivos específicos.

2.1 Reconocer los objetos especificados en el entorno que rodea al robot:

Se deben hacer diferentes pruebas para definir los parámetros y condiciones óptimas para realizar la captura de las imágenes. Después se utilizará Caffe (framework) el cual soporta muchos tipos diferentes de arquitecturas de aprendizaje profundo, en este caso se utilizará el proyecto *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks* [?], orientada a la clasificación de objetos en imágenes y la ubicación de la región donde se encuentre. Se utilizarán modelos de redes neuronales convolucionales pre-entrenados que se fundamenten en la base de datos ImageNet la cual está organizada de acuerdo a la jerarquía de sustantivos de WordNet en el que cada nodo de la jerarquía está representado por cientos y miles de imágenes, aunque también en caso de ser necesario se entrenará una red neuronal con un *dataset* enfocado en reconocer específicamente los objetos previamente detallados para obtener un modelo que presente mayores beneficios en comparación a los ya existentes. Los objetos específicos que debe reconocer el robot aún están por definir porque la mano en el brazo del robot no es muy grande y se deben hacer pruebas con objetos pequeños para definirlos. En pocas palabras, Caffe se encarga de recibir como **entrada** una imagen tomada por el robot y su **salida** es o son los sustantivos u objetos que más probabilidad tienen de coincidir con la realidad junto con la ubicación

de la región donde reconoció el objeto.

2.2 Reconocer la voz del paciente y traducirlo como una instrucción:

El reconocimiento de voz se llevará a cabo mediante bibliotecas que se importan en Python para cumplir este objetivo. Por ejemplo, el primer candidato para esta tarea es SpeechRecognition la cual es una biblioteca para Python soportado por la API Google Speech Recognition. Este tiene la función de convertir el mensaje del paciente a texto, y posteriormente el robot deberá interpretarlo como una instrucción. Se debe aclarar que después de la conversión de audio a texto, el algoritmo atenderá a instrucciones y objetos limitados, con el objetivo de simplificar el proyecto debido al corto periodo para desarrollarlo. Las palabras que debe reconocer según las métricas definidas son:

- Verbos: Traer, alcanzar.
- Sustantivos: Los tres nombres de los objetos que se deben localizar

2.3 Integrar las tareas de desplazamiento y el movimiento de objetos por parte del robot junto con el reconocimiento de objetos y de voz:

Para lograr la localización en el espacio del robot y el objeto deseado se hará uso del dispositivo tipo Kinect o sensor 3D para obtener información del entorno que rodea al robot y luego procesarla con PCL (Point Cloud Library) la cual incorpora utilidades para trabajar con nubes de puntos y procesamiento de imágenes 3D, esto tiene como finalidad de que el robot se ubique en el espacio que lo rodea pero también sirve para identificar dónde se encuentra el objeto deseado. Luego que se ha localizado el objeto, el robot debe desplazarse hasta las cercanías de este, tomar el objeto mediante un movimiento pregrabado (con el fin de evitar la ardua tarea de *Grasping* en robótica), regresar hasta donde se encuentra el paciente y colocarlo cerca según las especificaciones previamente definidas.

Para obtener toda esta información de los sensores del robot y también para accionar sus motores que permiten desplazarlo entre distintos puntos y mover sus extremidades se utilizará el sistema operativo para robots que se caracteriza por ser un software de tipo middleware (ROS: Robot Operating System), el cual opera de la siguiente manera: el núcleo donde se realiza todo el procesamiento de la información es el ordenador, el cual se conecta al robot que en este caso es una ramificación del núcleo mediante su IP dentro de una red, de esta manera el núcleo puede acceder a todos los dispositivos del robot como

cámara, sensores, kinect, micrófonos y también puede accionar los mecanismos para mover al robot y sus extremidades.

Algunos puntos a considerar:

- Delimitar la altura en la que se encuentra el objeto, la cual ronda un rango entre 70 cm a 100 cm
- La mano del robot o pinza se abrirá y cerrará en dirección paralela al suelo.
- El proyecto no tiene como requisito contemplar obstáculos.
- Este proyecto no tiene como requisito que su evaluación sea en un entorno real como un cuarto de habitación, un cuarto de hospital, entre otros; esto por el corto periodo para desarrollar el proyecto.

Por último se debe lograr enlazar cada tarea individual y crear un sistema funcional, en esta etapa es posible que se deban hacer modificaciones a las tareas individuales para mejorar los resultados del proyecto.

