

## Plan de Trabajo Final

### Carrera Ingeniería de Sistemas

#### Facultad de Ciencias Exactas – UNICEN

**Tema:** Framework para facilitar el reconocimiento de gestos en dispositivos móviles.

**Alumno/s:** Sr. Jorge Andrés Mangeruga

**Director:** Dr. Álvaro Soria

**Codirector:** Ing. Rodrigo Sebastian Ibañez

### 1. Motivación

El vertiginoso aumento en la capacidad de los dispositivos móviles y la variedad de sensores que ofrecen, posibilitan el desarrollo de novedosas aplicaciones en el campo del reconocimiento de actividades humanas. El uso de dispositivos móviles en la vida cotidiana de las personas ha crecido rápidamente [1], en consecuencia se encuentran presentes en todas las actividades que las personas realizan y el contexto en donde se mueven. Los sensores con los que ahora cuentan los dispositivos móviles logran transformar esa información en diferentes medidas cuantificables, tales como la aceleración de un movimiento del teléfono, la temperatura ambiente, o las posiciones de las partes del cuerpo al estilo Kinect. Gracias a la actual capacidad de cómputo se puede procesar esa información en vivo, por ejemplo, para realizar reconocimiento de gestos en aplicaciones tales como juegos, monitoreo de la salud, nuevas formas de interacción, robótica, entre otras [2].

Actualmente, existen algunos enfoques que ayudan a implementar este tipo de aplicaciones, pero que no logran ser una solución integral al problema. Los desarrolladores de aplicaciones que necesiten obtener datos de sensores de movimiento pueden aplicar los enfoques académicos mostrados en [1], [2], [3] y [4], pero deberían implementar por completo el sistema de reconocimiento. Esta problemática fue abordada por [5], un framework que permite entrenar y reconocer gestos con mucho menor esfuerzo por parte del desarrollador. Sin embargo, no soporta reconocimiento de gestos desde Kinect. Si algún desarrollador necesitara esta funcionalidad, el camino actual es la utilización de SDK's como MotionSDK<sup>1</sup> o eyeSight<sup>2</sup>. Ambos brindan una capa de abstracción para obtener información de alto nivel de las cámaras de los dispositivos, pero al igual que enfoques antes mencionados requieren que el desarrollador tenga conocimientos sobre las técnicas de reconocimiento existentes para poder aprovechar esa información.

En este contexto, la presente tesis propone proveer un framework para aplicaciones en dispositivos móviles que interactúen por medio de gestos, y simplifique la labor del desarrollador. El framework tomará como entrada información proveniente de un variado espectro de sensores o la cámara, y dará soporte a la definición y administración de gestos. Con respecto a la definición, el framework permitirá realizar entrenamientos para que la aplicación entienda cómo es cada gesto. Con respecto al reconocimiento, cuando el usuario de la aplicación final realice movimientos, el framework los comparará con los gestos entrenados y notificará a la aplicación para que reaccione en consecuencia. Para dar soporte a estas características, se explorará el concepto de machine learning, que por medio de distintas técnicas de reconocimiento de gestos dotará de inteligencia al framework.

<sup>1</sup> <http://www.bitgym.com/developer/motion-sdk>

<sup>2</sup> <http://eyesight-tech.com/>

## 2. Propuesta de trabajo

Este trabajo tiene como objetivo el desarrollo de un framework, cuyo nombre es **DronGR**, que asista al desarrollador en la incorporación de reconocimiento de gestos en una aplicación para dispositivos móviles. En este contexto, un gesto se define como la secuencia de las posiciones de una parte del cuerpo o los datos obtenidos de alguno de los sensores del móvil, a lo largo del tiempo. El framework está basado en un enfoque de **machine learning**, que combinará diferentes técnicas de inteligencia artificial para aprender y luego reconocer gestos. En la *Figura 1* se muestra una vista conceptual general del enfoque propuesto.

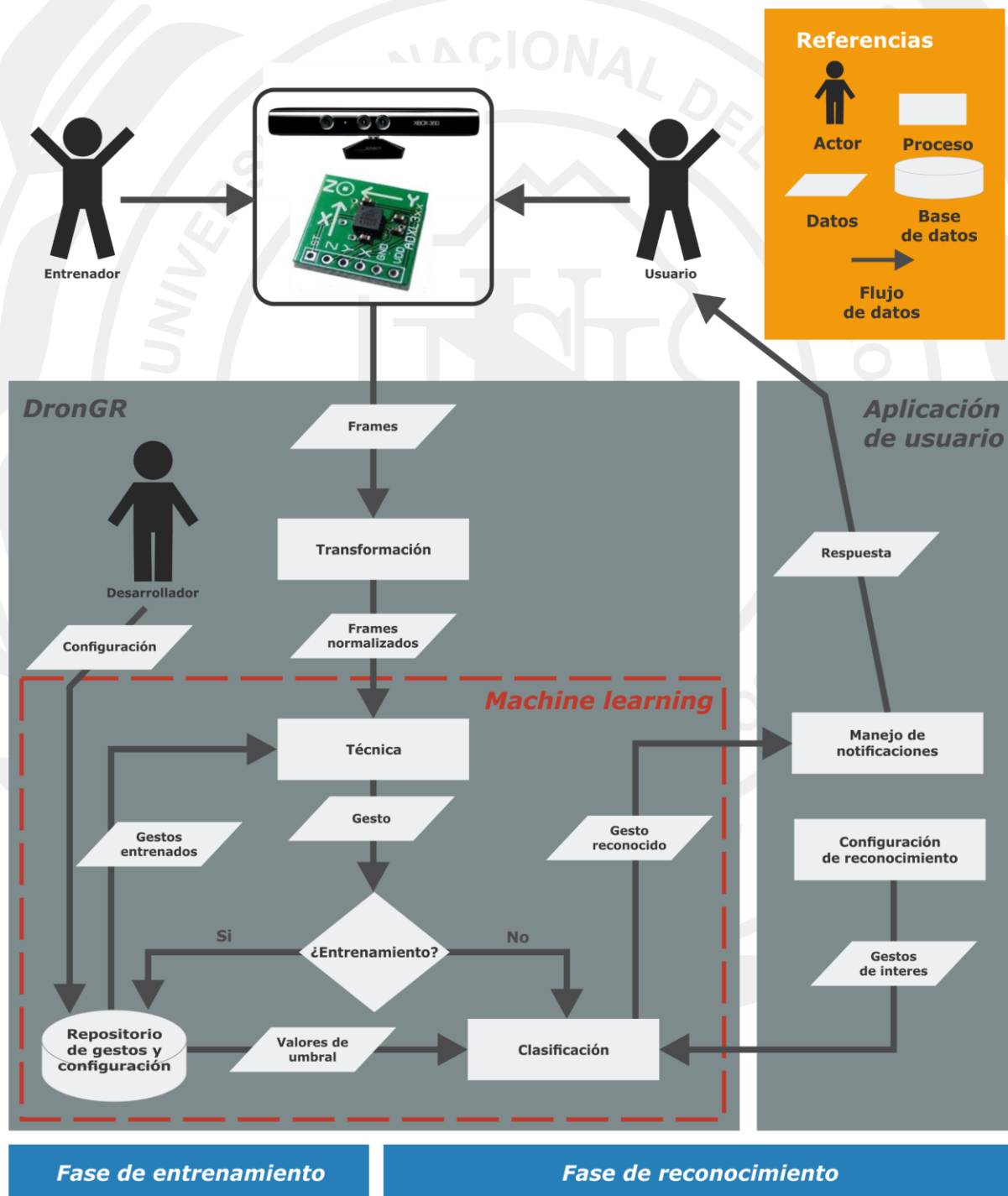


Figura 1: Vista conceptual de DronGR

El enfoque tiene dos partes bien diferenciadas, la **fase de entrenamiento** y la **fase de reconocimiento**. Durante la fase de entrenamiento, el **entrenador** realiza un gesto, el cual es capturado por la cámara o alguno de los sensores del dispositivo móvil. Dado que los datos en crudo son de diferentes tipos dependiendo del sensor utilizado, es necesario aplicar una **transformación** que tendrá como objetivo homogeneizarlos, de forma que el framework trabaje siempre con datos equivalentes. Una vez transformados, los datos son dados de entrada a la **técnica** de reconocimiento, junto con la información previa que haya sobre los gestos, para ser procesados y aumentar el conocimiento que se tiene sobre los mismos.

Entre las técnicas que serán analizadas por este trabajo se encuentran: HMM (*Hidden Markov Model*), son métodos estocásticos para modelos temporales y secuencia de datos [6]; DTW (*Dynamic Time Warping*), es un algoritmo utilizado para medir las similitudes entre dos secuencias de datos las cuales pueden variar en tiempo o velocidad [7]; Análisis de Procrustes, es un proceso de superposición de objetos llevado a cabo mediante la óptima traslación, rotación y escalado uniforme [8]; y String Matching, son algoritmos que intentan encontrar un patrón o cadena de caracteres dentro de un texto o dentro de otra cadena más larga [9].

Luego de procesar los datos, el resultado de la técnica es almacenado en el **repositorio de gestos y configuración**, junto a la configuración particular que requiera ese gesto. Este proceso genera información más precisa acerca de cómo es un gesto. Cuanta mayor sea la cantidad de veces que este proceso se repita, el conocimiento que se tenga sobre cada gesto será mejor, lo que repercutirá en una mayor tasa de reconocimiento.

La fase de reconocimiento, es iniciada desde la **aplicación de usuario** construida sobre el framework. Este recibe proveniente de la aplicación, cuáles gestos se desean contemplar durante el reconocimiento. De esta manera, los gestos realizados por el **usuario** también son capturados y transformados, para ser dados de entrada a la **técnica** que compara el conocimiento existente sobre el gesto con los datos del nuevo gesto realizado por el usuario. A continuación se realiza la **clasificación** del gesto, es decir, la información producida por la técnica y los valores de umbral se utilizan para decidir cuál fue el gesto realizado. Una vez determinado esto, el framework notifica a la aplicación para ejecutar una acción en respuesta que sea vista por el usuario.

En resumen, el enfoque propuesto le facilitará al desarrollador la tarea de reconocer gestos usando dispositivos móviles. El empleo del framework DronGR permitirá grabar, entrenar y definir gestos por medio de una interfaz gráfica amigable. Mediante archivos de configuración se podrán configurar los gestos y las respectivas acciones que deberá realizar la aplicación en respuesta a dicho gesto. Las técnicas utilizadas ofrecerán niveles aceptables de reconocimiento, lo que asegurará ser un framework confiable para ser integrado en aplicaciones de terceros, permitiendo a los desarrolladores focalizar sus esfuerzos en implementar la interfaz adecuada para el usuario.

### 3. Cronograma de actividades

En la tabla se muestra la planificación realizada para las tareas de incumbencia en este proyecto. Su extensión alcanza los 12 meses y medio. Vale aclarar, que ya se han logrado avances para algunas de ellas.

N° Actividad	Descripción	Duración
1	Relevamiento bibliográfico de los enfoques actuales	2 meses
2	Planteo y análisis de los requerimientos del problema	1 mes
3	Implementación del framework	4 meses
4(*)	Elaboración de experimentos y análisis de resultados	2 meses
5(*)	Redacción del informe final	5 meses
6	Preparación de la defensa de tesis	½ mes

(\*) Actividades en paralelo.

#### 4. Bibliografía

- [1] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," in *Ambient Assisted Living and Home Care*, Springer, 2012, pp. 216–223.
- [2] M. F. A. bin Abdullah, A. F. P. Negara, M. S. Sayeed, D.-J. Choi, and K. S. Muthu, "Classification algorithms in human activity recognition using smartphones," *World Acad. Sci. Eng. Technol.*, vol. 68, pp. 422–430, 2012.
- [3] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explor. Newsl.*, vol. 12, no. 2, pp. 74–82, 2011.
- [4] A. Rasekh, C.-A. Chen, and Y. Lu, "Human activity recognition using smartphone," *Tex. AM Univ Tech Rep*, 2011.
- [5] J. Alexandersson, "Ein Toolkit zur automatischen Klassifikation von Gesten," *Masters Thesis DFKI*, 2008.
- [6] C.-Y. Kao and C.-S. Fahn, "A human-machine interaction technique: hand gesture recognition based on hidden Markov models with trajectory of hand motion," *Procedia Eng.*, vol. 15, pp. 3739–3743, 2011.
- [7] S. Salvador and P. Chan, "Toward accurate dynamic time warping in linear time and space," *Intell. Data Anal.*, vol. 11, no. 5, pp. 561–580, 2007.
- [8] A. Ross, "Procrustes analysis," *Course Rep. Dep. Comput. Sci. Eng. Univ. S. C.*, 2004.
- [9] P. D. Michailidis and K. G. Margaritis, "On-line string matching algorithms: Survey and experimental results," *Int. J. Comput. Math.*, vol. 76, no. 4, pp. 411–434, 2001.