

20, 40, 60 y 80 muestras de cada gesto. El objetivo de esta prueba es determinar la cantidad mínima de muestras necesarias para obtener una precisión aceptable, y conocer hasta qué punto las técnicas incrementan su precisión al incrementar la cantidad de muestras utilizadas para entrenamiento.

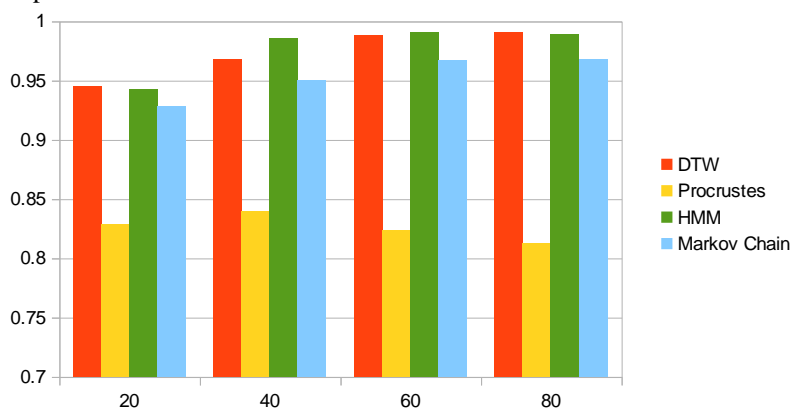


Fig. 3. Grafico que compara la cantidad de gestos reconocidos al variar la cantidad de muestras usadas para el entrenamiento

La Figura 3 muestra los resultados obtenidos. Los tamaños de dataset utilizados se listan en el eje horizontal mientras que la precisión alcanzada se lista en el eje vertical. Del gráfico se puede observar que incrementar la cantidad de muestras utilizadas para el entrenamiento mejora la precisión de DTW un 4.1% cuando se incrementa de 20 a 80 el tamaño del dataset de prueba. Si bien este aumento es notorio, se ve desacelerado al llegar a un dataset de 40 muestras por gesto. Por ejemplo, HMM mejora la precisión en 0.35% cuando se incrementa de 40 a 80 el tamaño del dataset.

4 Conclusiones

En este trabajo hemos evaluado las técnicas *Dynamic Time Warping*, *Procrustes Analysis*, *Markov Chain* y *Hidden Markov Models* en el reconocimiento de gestos corporales, las cuales utilizan las posiciones de las partes del cuerpo provistas por Kinect. Los experimentos mostraron que *Dynamic Time Warping* y *Hidden Markov Models* son las más precisas reconociendo 99,1% y 98,9% de los gestos respectivamente. La técnica con menor precisión fue *Procrustes Analysis* reconociendo el 81,2% de los gestos. Además, como era de esperarse, el incremento de la cantidad de muestras utilizadas para entrenamiento incrementa la precisión de las técnicas. Sin embargo, este incremento no es tan significativo al superar las 40 muestras por gesto.

Como trabajo futuro sería importante medir el tiempo requerido por cada técnica tanto para ser entrenada como para evaluar un nuevo gesto. Además, ver cuánto influye en este tiempo la cantidad de gestos a reconocer en simultáneo. De esta forma podríamos saber cuál se ajusta mejor para ser utilizada en tiempo real.

Por último, sería interesante realizar nuevas pruebas con un conjunto más variado de gestos, incorporando otras técnicas de inteligencia artificial como *Support Vector Machines* y *Decision Trees*.

Agradecimientos

Agradecemos el apoyo financiero brindado por la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) a través del proyecto PICT 2011-0080.

Bibliografía

- [1] J. Boehm, “Natural user interface sensors for human body measurement,” *Int Arch Photogramm Remote Sens Spat. Inf Sci*, vol. 39, p. B3, 2012.
- [2] E. A. Suma, B. Lange, A. Rizzo, D. M. Krum, and M. Bolas, “FAAST: The Flexible Action and Articulated Skeleton Toolkit,” in *Virtual Reality Conference (VR), 2011 IEEE*, 2011, pp. 247–248.
- [3] F. Kistler, B. Endrass, I. Damian, C. Dang, and E. André, “Natural interaction with culturally adaptive virtual characters,” *J. Multimodal User Interfaces*, vol. 6, no. 1–2, pp. 39–47, 2012.
- [4] V. Thiruvarduchelvan and T. Bossomaier, “Towards realtime stance classification by spiking neural network,” in *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 2012, pp. 1–8.
- [5] A. Bleiweiss, D. Eshar, G. Kutliroff, A. Lerner, Y. Oshrat, and Y. Yanai, “Enhanced interactive gaming by blending full-body tracking and gesture animation,” in *ACM SIGGRAPH ASIA 2010 Sketches*, 2010, pp. 34:1–34:2.
- [6] X. Yang and Y. Tian, “EigenJoints-based action recognition using Naïve-Bayes-Nearest-Neighbor,” in *CVPR Workshops*, 2012, pp. 14–19.
- [7] S. Bhattacharya, B. Czejdo, and N. Perez, “Gesture classification with machine learning using kinect sensor data,” in *Emerging Applications of Information Technology (EAIT), 2012 Third International Conference on*, 2012, pp. 348–351.
- [8] C. Waithayanon and C. Aporntewan, “A Motion Classifier for Microsoft Kinect,” in *Computer Sciences and Convergence Information Technology (ICCIT), 2011 6th International Conference on*, 2011, pp. 727–731.
- [9] S. Salvatore and P. Chan, “FastDTW: Toward accurate dynamic time warping in linear time and space,” presented at the 3rd Workshop on Mining Temporal and Sequential Data, 2004.
- [10] A. Ross, “Procrustes analysis,” *Course Rep. Dep. Comput. Sci. Eng. Univ. S. C.*, 2004.
- [11] K. Lange, “Finite-State Markov Chains,” in *Numerical Analysis for Statisticians*, Springer, 2010, pp. 503–526.
- [12] L. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [13] K. Lai, J. Konrad, and P. Ishwar, “A gesture-driven computer interface using Kinect,” in *Image Analysis and Interpretation (SSIAI), 2012 IEEE Southwest Symposium on*, 2012, pp. 185–188.