

REAL-TIME MARKERLESS MOTION CAPTURE

OBJETIVOS:

Crear una aplicación open-source que realice captura de movimiento facial sin la necesidad de usar marcadores ni entornos controlados (“in the wild”) y sea capaz de procesar video capturado en tiempo real (al menos).



Figura 1. Proceso seguido por la aplicación

RESULTADOS:

- La aplicación se desarrollo utilizando las bibliotecas open-source de OpenCV y OpenGL, así como el framework Qt para la interfaz gráfica.
- Los parámetros de los algoritmos se ajustaron para lograr realizar la captura en tiempo real, el paso que más tiempo consume es la detección facial y no la localización de hitos.

CONCLUSIONES:

- La falta de equipo caro como cámaras o renta de estudios se puede compensar con poder computacional hasta cierto punto.

REFERENCIAS:

1. Chen Cao, Derek Bradley, Kun Zhou, and Thabo Beeler. Real-time high-fidelity facial performance capture. ACM Transactions on Graphics (ToG), 34(4):46, 2015.
2. Vahid Kazemi and Josephine Sullivan. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
3. P. Garrido D. Bradley T. Beeler P. Pérez M. Stamminger M. Nießner C. Theobalt M. Zollhofer, J. Thies. State of the art on monocular 3d face reconstruction, tracking, and applications. Eurographics, 37(2), 2018.
4. Alberto Menache. Understanding Motion Capture for Computer Animation. Elsevier, 2011.
5. Thibaut Weise, Sofien Bouaziz, Hao Li, and Mark Pauly. Realtime performance-based facial animation. In ACM transactions on graphics (TOG), volume 30, page 77.ACM, 2011.

DISEÑO DE LA SOLUCIÓN:

- Proceso de tres partes, captura, retargeting y animación (ver Figura 1).
- Uso de las implementaciones de cascadas de Haar y el algoritmo de V. Kazemi[2] incluidas en OpenCV para las tareas de detección facial y reconocimiento de hitos faciales (facial landmarks) respectivamente como se muestra en la Figura 2.
- Generación de animación mediante OpenGL
- La restricción de “tiempo real” se considero como 24 cuadros por segundo (0.041 segundos) al ser la razón de captura comúnmente utilizada para video

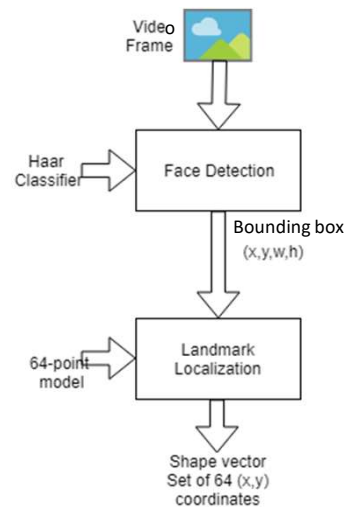


Figura 2. Proceso de detección y localización de hitos faciales