

Desarrollo de un banco de imágenes de rostros mediante la adquisición de muestras utilizando una cámara ToF

1st Rafael Duarte-Pérez
Facultad de Informática (FIF)
Universidad Autónoma de Querétaro (UAQ)
Querétaro, México
rafael.duarte@uaq.mx

2th Gabriela Pacheco-Sánchez
Facultad de Informática (FIF)
Universidad Autónoma de Querétaro (UAQ)
Querétaro, México
gabriela.pacheco@uaq.mx

3th Martín Muñoz-Mandujano
Facultad de Informática (FIF)
Universidad Autónoma de Querétaro (UAQ)
Querétaro, México
martin.munoz.mandujano@uaq.mx

4th Jorge-Luis Pérez-Ramos
Facultad de Informática (FIF)
Universidad Autónoma de Querétaro (UAQ)
Querétaro, México
jorgelupr@gmail.com

5th Luis-Antonio Díaz-Jiménez
Facultad de Informática (FIF)
Universidad Autónoma de Querétaro (UAQ)
Querétaro, México
luis.diaz@uaq.mx

6th Hugo Jiménez-Hernández
Facultad de Informática (FIF)
Universidad Autónoma de Querétaro (UAQ)
Querétaro, México
hugo.jimenez@uaq.mx

Resumen—El reconocimiento facial mediante técnicas de visión por computadora, se ha vuelto una tecnología cada vez más relevante en diversas aplicaciones. Si bien existen sistemas que emplean imágenes a color, la incorporación de información de profundidad puede enriquecer significativamente estos modelos. Sin embargo, la escasez de conjuntos de datos que incluyan tanto imágenes a color como a profundidad representa un desafío en este campo. Este trabajo presenta una metodología para la creación de una Base de Datos de rostros, utilizando un grupo de personas de la Facultad de Informática de la Universidad Autónoma de Querétaro. La Base de Datos se conforma de 54 individuos. De cada individuo se tiene un total 540 imágenes en formatos: color y profundidad, con diferentes perspectiva del rostro. Para la adquisición de estos datos se utilizó una cámara Intel Depth Camera D415, que emplea tecnología de tiempo de vuelo (ToF).

Palabras clave—Base de datos, imágenes a profundidad, cámara TOF

I. INTRODUCCIÓN

I-A. Antecedentes

La identificación de rostros que se obtiene a través de la visión por computadora se ha estudiado y trabajado a lo largo de los años dando pie a múltiples aplicaciones como el entretenimiento, marketing, salud y especialmente en el entorno de los sistemas de seguridad [4]. Durante las etapas involucradas en la detección de rostros, es crítico que se tenga una base de información de rostros que sirva para probar, desarrollar

y comprobar el reconocimiento facial [3]. En la actualidad existen diferentes bases de datos que proveen imágenes de rostros que son usadas para desarrollar, probar y validar algoritmos de reconocimiento facial [5]. A continuación se enlistan algunos ejemplos de datasets relevantes en el estado del arte:

- Olivetti Research Laboratory (ORL): Desarrollada por el Olivetti Research Laboratory, incluye 400 imágenes de 40 individuos diferentes. Las fotografías fueron capturadas en distintos momentos y presentan variaciones de iluminación, expresiones faciales diversas y posibles oclusiones en los rostros [8].



Figura 1. Muestra de las imágenes de ORL

- Institute of Automation, Chinese Academy of Sciences, Near Infrared – Visible Light (CASIA NIR-VIS): Base

de datos que contiene imágenes de los rostros de 725 personas diferentes con dos tipos de formato diferentes; JPEG con un número de 1-22 imágenes por persona y BMP con un número de 5-50 imágenes por persona. Para ambos casos se utilizó una resolución de 640x480 píxeles [11].



Figura 2. Muestra de las imágenes de CASIA NIR-VIS

- RGB Depth and Thermal (RGB-D-T): Base de datos con 45,900 imágenes de 51 personas diferentes, orientada a hombres entre 20 - 40 años, el sensor utilizado fue un Microsoft Kinect y capturó 3 tipos de imágenes; RGB con dimensiones de 640x480 píxeles, imágenes de profundidad con dimensiones de 640x480 píxeles e imágenes térmicas con una resolución de 384x288 píxeles [10].



Figura 3. Muestra de las imágenes de RGB-D-T

A partir de la revisión del estado del arte, se expone la Tabla 1 con la finalidad de comparar algunas bases de datos relacionadas con los rostros tanto a color como a profundidad [2],[6],[7],[8],[10],[11] y [12].

Se utiliza la Fig. 4 de [1] con la finalidad de ilustrar los aportes en una línea del tiempo.

Tabla I
COMPARACIÓN DE BASES DE DATOS DE IMÁGENES DE ROSTROS

Base de datos	Características	Nº de sujetos	Nº de imágenes
ORL	40 personas, variaciones de iluminación y expresiones.	40	400
FERET	Variaciones de poses, expresiones.	1,199	14,051
AR	Variaciones de iluminación, expresiones, oclusiones.	126	4,000+
XM2VTS	Secuencias de video, variaciones de iluminación.	295	2,360+
BANCA	Expresiones faciales en ambientes controlados/degradados.	52	5,280
FRGC	Variabilidad en iluminación y expresiones.	466	50,000
LFW	Ambientes no controlados, variaciones de poses.	5,749	13,000+
CMU Multi-PIE	Variaciones de iluminación, poses, expresiones.	337	750,000+
CASIA NIR-VIS	Imágenes en infrarrojo y visible.	725	17,580
IIB-A	Iluminación y poses variadas.	500	13,000
MF	Imágenes de internet, desafío de identificación masiva.	690,000	1M+
VGG Face2	Variaciones en iluminación, edad, poses.	9,131	3.31M
DFW	Rostros disfrazados en ambientes no controlados.	1,000	11,000+
RGB-D-T	Imágenes RGB, de profundidad y térmicas.	51	30,000+
RMFRD	Rostros manipulados.	90,000	200,000+

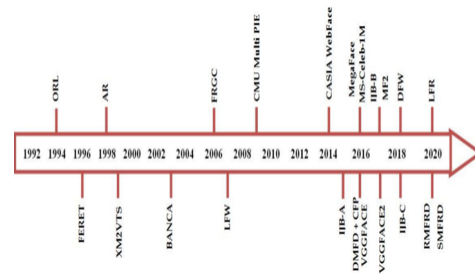


Figura 4. Línea de tiempo de las bases de datos de rostros

De la revisión de la literatura se identificaron áreas de oportunidad, especialmente desde el perspectiva del acceso a muestras y la manera en que se representa la imagen a profundidad.

I-B. Motivación

El uso de imágenes de profundidad tiene el objetivo de proveer información sobre la topología del rostro para lograr un reconocimiento facial resistente a las limitaciones como lo son la iluminación, la oclusión y la perspectiva [1]. Según [9] la precisión del reconocimiento facial con imágenes a color es insuficiente debido a su sensibilidad a factores como la iluminación, la pose y las variaciones en las expresiones. Sin embargo, el uso de imágenes a profundidad puede mitigar en gran medida estas limitaciones. El número de métodos que

incorporan imágenes a profundidad para el reconocimiento facial está en aumento. [12] indica que es necesario desarrollar modelos de reconocimiento facial que consideren la extracción de características a partir de imágenes de profundidad, para estos modelos debe existir un banco de imágenes que sirva para realizar el proceso de entrenamiento o validación.

Este trabajo surge para proveer un banco de imágenes a profundidad que, a diferencia de [10], provea la información de distancia en la imagen ya que en el dataset de [10] sólo se trabajan con archivos PNG.

I-C. Objetivos

Desarrollar un banco de imágenes de rostros tanto a color como a profundidad, cuyas características sean adecuadas para su uso en modelos de reconocimiento facial.

Dar acceso abierto al banco de imágenes a partir de la aprobación del H. Consejo Facultativo de la Facultad de Informática de la Universidad Autónoma de Querétaro.

II. METODOLOGÍA

II-A. Información de los individuos

Las capturas de los rostros contenidas en este banco de imágenes fueron adquiridas gracias a la participación de estudiantes en las licenciaturas ofertadas por la Facultad de Informática de la Universidad Autónoma de Querétaro (UAQ). Se cuenta con el consentimiento de los estudiantes voluntarios. Las edades de la muestra corresponden entre 18 a 24 años de edad, donde predomina la presencia del sexo masculino debido a la disponibilidad de muestras. Por cuestiones de privacidad no se guardan nombres, expedientes u alguna otra información de los rostros dentro del banco de imágenes.

II-B. Propiedades del entorno de captura

La captura se realizó con las siguientes consideraciones:

- Pantalla verde, para el fondo.
- Un par de sombrillas fotográficas, con la finalidad de obtener la misma iluminación en las capturas a color.
- Tripié, para obtener una altura homogénea entre las muestras.
- Silla, para dar comodidad al individuo durante la captura de muestras.

A lo largo del proceso de captura, se encontraron situaciones donde se tuvieron que hacer ajustes al tripié ya que los individuos variaban en sus estaturas. Estos cambios eran principalmente de altura y ángulo vertical de la toma, sin embargo, se regresaba a los valores originales después de las tomas.

La distribución de los elementos se realizó a partir de ajustes empíricos en capturas de prueba, este proceso resultó en el arreglo de elementos que se muestra en la Fig. 5.

El objetivo es centrado en la pantalla verde. El color azul hace referencia a la distancia del fondo a la cámara y de la silla a la cámara. El color rojo hace referencia a la altura de la silla con respecto al suelo y la altura de la cámara con respecto del suelo.

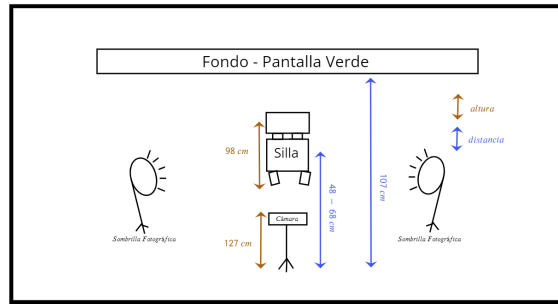


Figura 5. Medición y distribución del entorno de captura.

II-C. Propiedades de las capturas adquiridas

Para la captura de muestras se utilizó una cámara RGB-Depth, se utilizó este sensor por la captura simultánea de muestras a color y a profundidad que provee. Podemos observar una muestra de la captura simultánea en la Fig. 6. Este banco de imágenes se construyó con el objetivo de trabajar imágenes a profundidad, sin embargo, se guardaron las imágenes a color para su uso en trabajos futuros.

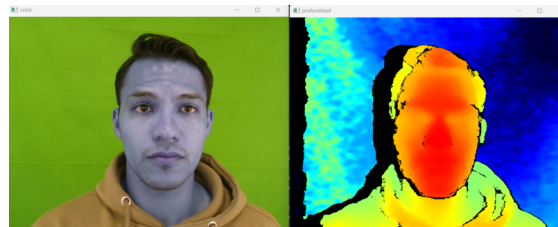


Figura 6. Captura a color y a profundidad, de izquierda a derecha respectivamente

Las muestras obtenidas tienen una resolución de 640 x 480 píxeles. Los formatos elegidos fueron JPG para las imágenes a color y archivo de texto (.txt) para las imágenes a profundidad. En la Fig. 7 se puede observar la muestra a profundidad graficada y el archivo de texto correspondiente.

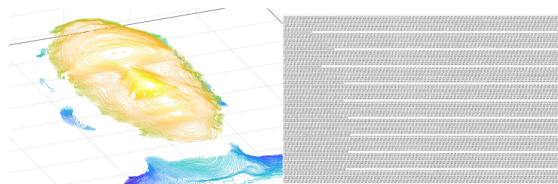


Figura 7. Muestra de profundidad graficada y su formato original

El archivo de textos sirvió para poder representar las filas y columnas de píxeles, mientras que a su vez, el valor de cada pixel representaba la distancia medida por el sensor.

Las muestras de este banco de imágenes contienen rostros con 3 expresiones: sonriendo, neutral y con ceño fruncido. Se buscó que se tuviera una variedad en las perspectivas del

rostro, por lo que se recolectaron 9 perspectivas diferentes a cada individuo (véase la Fig. 8 y Fig. 9):

- Perfil de frente
- Perfil de frente con inclinación a la derecha
- Perfil de frente con inclinación a la izquierda
- Perfil girando hacia arriba
- Perfil girando hacia arriba con inclinación a la derecha
- Perfil girando hacia arriba con inclinación a la izquierda
- Perfil girando hacia abajo
- Perfil girando hacia abajo con inclinación a la derecha
- Perfil girando hacia abajo con inclinación a la izquierda

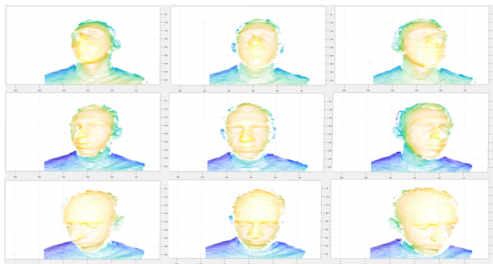


Figura 8. Perspectivas de profundidad capturadas en el banco de imágenes



Figura 9. Perspectivas a color capturadas en el banco de imágenes

En la búsqueda de tener una cantidad robusta de datos, previendo tomas en movimiento y fallas del sensor a distancia, se tomaron 9 capturas de cada perspectiva, dichas capturas están separadas una de otra por 0.2 segundos de diferencia.

II-D. Organización de las muestras

El proceso de almacenamiento de las muestras se diseñó con la finalidad en que todas las muestras de un individuo, tanto a color como a distancia, estuvieran concentradas en un mismo lugar, pero que a su vez estuvieran separadas por cada categoría.

Para cumplir con lo anterior, se crearon subcarpetas dentro de la carpeta de cada individuo, donde los archivos poseen nombres similares pero la diferencia radica en la carpeta donde

están designados. En la Fig. 10 se muestra un resumen del patrón para alojar las muestras de los individuos.

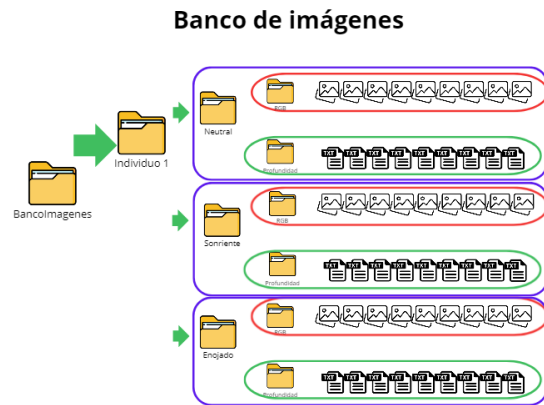


Figura 10. Organización del almacenamiento de las muestras por individuo

El tamaño de la BD es de aproximadamente 60 Giga Bytes contando las muestras a color y a profundidad que se han capturado de los 60 individuos.

II-E. Proceso de adquisición de muestras

En condiciones controladas, se procedió con la captura de muestras, este proceso registró tiempos de entre 30 y 40 minutos para la adquisición de un solo individuo en su primer diseño.

Se desarrolló un sistema que permitió automatizar las capturas de muestras con la finalidad de evitar la pérdida de tiempo en la adquisición y guardado manual de las muestras. Dicho sistema se denominó Sistema Capturador de Rostros a Color y a Profundidad (SCRCP). El proceso de adquisición de muestras redujo considerablemente su tiempo de adquisición, tomando aproximadamente entre 5 y 8 minutos por muestra.

Se estandarizó el proceso de captura de acuerdo al algoritmo presentado en la Fig. 11:

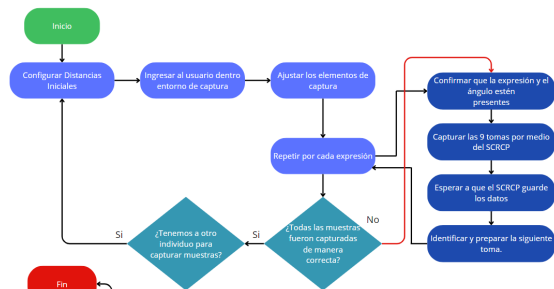


Figura 11. Proceso de captura de imágenes

III. RESULTADOS, DISCUSIÓN Y CONCLUSIONES

El producto de este proyecto fue un banco de imágenes enfocado en rostros de personas entre los 18 y 24 años de edad,

donde las muestras son mayormente del género masculino. Las muestras obtenidas son homogéneas en formato, entorno, poses y expresiones.

Este banco de imágenes posee muestras de los rostros de 60 individuos tanto a color como a profundidad, para cada formato se capturó una combinación de 3 expresiones con 9 perspectivas diferentes, cada captura cuenta con 9 imágenes separadas por 0.2 segundos. Por lo anterior, el banco de imágenes cuenta con un total 29,160 muestras de rostros, organizadas en carpetas según el individuo, expresión y perspectiva capturada (véase la Tabla II).

Tabla II
CARACTERÍSTICAS DEL BANCO DE IMÁGENES

Número de poses capturadas	9 poses: De frente, frente derecha, frente izquierda, arriba frente, arriba izquierda, arriba derecha, abajo frente, abajo derecha y abajo izquierda
Número de expresiones capturadas	3 expresiones: Neutral, sonriente y ceño fruncido
Número capturas por pose	18 imágenes por pose: 9 a profundidad y 9 a color. Cada captura separada por 0.2 segundos
Número total de capturas por individuo	486: 18 fotos por pose, cada pose tiene 3 expresiones
Resolución de las imágenes en el banco de imágenes	Ambos formatos comparten la misma resolución, 640 x 480 píxeles

Adicional, se desarrolló un sistema en lenguaje Python con la finalidad de automatizar el proceso de adquisición de muestras. El resultado fue un sistema que permite:

1. Visualizar la lectura a color y a profundidad a fin de preparar la captura.
2. Elegir el momento de captura.
3. Realizar el guardado de las muestras en las carpetas correspondientes.
4. Realizar el proceso de captura de manera iterativa con el mínimo esfuerzo del usuario.

El propósito de uso para este banco de imágenes requirió de una distancia significativa del rostro con el fondo de la imagen, lo suficiente para hacer posible su exclusión. Se realizó de esta manera para que en el preprocesamiento de las imágenes a color y a profundidad simple retirar el fondo de las muestras.

IV. CONCLUSIONES

En este trabajo se expone el proceso que se realizó para poder crear un banco de imágenes tanto a color como a profundidad de rostros. El banco de imágenes creado resultó útil en su uso para un proyecto de investigación relacionado al reconocimiento facial con imágenes a profundidad. Las características de las muestras facilitaron el preprocesamiento de las imágenes con la finalidad de excluir información ajena al rostro del individuo.

La creación del banco de imágenes surgió a partir de no poder encontrar un banco de imágenes a profundidad libre y con las características deseadas, por lo que un propósito adicional del presente trabajo es contribuir al estado del arte con un banco de imágenes abierto que sea utilizado por los

interesados bajo aceptación del H. Consejo Académico de la Facultad de Informática de la UAQ. Las solicitudes se canalizarán por medio del correo de rafael.duarte@uaq.mx.

Como trabajo futuro se pretende continuar con la adquisición de muestras de individuos con la finalidad de incrementar las muestras en el banco de imágenes, así mismo se plantea la creación de un nuevo banco de imágenes que contemple muestras con ruidos, perspectivas atípicas y oclusiones del rostro.

V. REFERENCIAS

REFERENCIAS

- [1] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, Present, and Future of Face Recognition: A Review," *Electronics*, vol. 9, no. 8, p. 1188, 2020. DOI: 10.3390/electronics9081188.
- [2] R. Gross, I. Matthews, J. F. Cohn, T. Kanade, and S. Baker, "Multi-PIE," in *Proceedings of the Eighth IEEE International Conference on Automatic Face and Gesture Recognition*, 2008.
- [3] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face Recognition Systems: A Survey," *Sensors*, vol. 20, no. 2, p. 342, 2020. MDPI AG. Available: <http://dx.doi.org/10.3390/s20020342>.
- [4] C. A. Madrigal-González, J. L. Ramírez-Madrigal, J. C. Hoyos-Arbeláez, and D. Stephen-Fernández, "Diseño de un sistema biométrico de identificación usando sensores capacitivos para huellas dactilares," *Revista Facultad de Ingeniería Universidad de Antioquia*, pp. 21-32, 2007.
- [5] D. Mahouachi and M. A. Akhloufi, "Recent Advances in Infrared Face Analysis and Recognition with Deep Learning," *AI*, vol. 4, no. 1, pp. 199-233, 2023. MDPI AG. Available: <http://dx.doi.org/10.3390/ai4010009>.
- [6] J. Matas, M. Hamouz, K. Jonsson, J. Kittler, Y. Li, C. Kotropoulos, A. Tefas, I. Pitas, T. Tan, H. Yan, F. Smeraldi, J. Bigun, N. Capdevielle, W. Gerstner, S. Ben-yacoub, Y. Abdeljaoued, and N. Mayoraz, "Comparison of face verification results on the XM2VTS database," 2000.
- [7] K. Messer et al., "Face authentication test on the BANCA database," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, Cambridge, UK, 2004, pp. 523-532. Vol. 4, doi: 10.1109/ICPR.2004.1333826.
- [8] Olivetti Research Laboratory, *The ORL database of faces*. Cambridge, UK: ATT Laboratories, 1994.
- [9] A. Ouamane et al., "Robust multimodal 2D and 3D face authentication using local feature fusion," *Signal, Image and Video Processing*, vol. 10, no. 1, pp. 129-137, 2014. Available: <https://doi.org/10.1007/s11760-014-0712-x>.
- [10] M. O. Simón, C. Corneanu, K. Nasrollahi, O. Nikisins, S. Escalera, Y. Sun, H. Li, Z. Sun, T. B. Moeslund, and M. Greitans, "Improved RGB-D-T based face recognition," *IET Biometrics*, vol. 5, no. 4, pp. 297-303, 2016.
- [11] L. Z. Stan et al., "The CASIA NIR-VIS 2.0 Face Database," *Computer Vision Foundation*, 2013.
- [12] S. Soltanpour, B. Boufama, and Q. M. Jonathan Wu, "A survey of local feature methods for 3D face recognition," *Pattern Recognition*, vol. 72, pp. 391-406, 2017. Available: <https://doi.org/10.1016/j.patcog.2017.08.003>.
- [13] P. J. Phillips et al., "Overview of the face recognition grand challenge," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005, pp. 947-954. vol. 1, doi: 10.1109/CVPR.2005.268.