

**Reconocimiento de Rostros**  
**Una solución orientada a modelos numéricos 3D**

**GRANATO RUBEN OMAR**  
Facultad de Cs. Exactas  
Univ. Nac. del Centro Prov. Bs. As.  
[rgranato@hotmail.com](mailto:rgranato@hotmail.com)

**TRIGO ARANA FRANCISCO JAVIER**  
Facultad de Cs. Exactas  
Univ. Nac. del Centro Prov. Bs. As.  
[fco\\_trigo@yahoo.com](mailto:fco_trigo@yahoo.com)

**ZELASCO JOSE FRANCISCO**  
Univ. de Buenos Aires - Facultad de Ingeniería  
Univ. Nac. del Centro Prov. Bs. As. - Facultad de Cs. Exactas  
[mailto:jfz@fecic.edu.ar](mailto:mailto:jfz@fecic.edu.ar)

**Resumen**

Este trabajo se orienta al reconocimiento de caras basado en la obtención de los modelos numéricos 3D de las mismas para su posterior comparación.

Para ello es necesario disponer de pares estereoscópicos de imágenes de caras con las que se obtendrán los modelos numéricos. Estos se obtienen utilizando métodos de puesta en correspondencia de puntos homólogos en estereología.

Luego de un procedimiento de normalización se comparan los modelos numéricos a los efectos del reconocimiento, para ello se hace referencia a métodos de evaluación de calidad de modelos numéricos.

Estos métodos podrían ser útiles, ya que los modelos numéricos son más invariantes que las imágenes monoculares de las caras. Se pueden integrar con los métodos de reconocimiento, usando imágenes monoculares y mejorar integrando otras aplicaciones que usen modelos aún mas sofisticados de respuesta perceptual en visión humana.

**Palabras Claves:** *Estereología, modelos numéricos, reconocimiento, visión computarizada.*

## **1.-Introducción**

La posibilidad de integración de muchos factores de identificación tales como password, tarjetas de identificación, voz, huellas digitales, rostros y otros robustecen la seguridad de tales sistemas en caso de ser usados por un conjunto de personas seleccionadas.

El reconocimiento de las caras se efectúa utilizando imágenes monoculares [YUI89]. En este trabajo se implementaron estrategias de restitución automática de pares estereoscópicos y reconocimiento de modelos numéricos de superficies de rostros.

Este estudio se orienta a la aplicación de algoritmos en el tratamiento de imágenes estereoscópicas que permiten optimizar el problema de puesta en correspondencia de puntos homólogos de dos o más imágenes de rostros aplicado al reconocimiento de personas, con el objetivo de lograr un modelo numérico adecuado de su superficie.

Existe una amplia literatura que propone distintos tipos de algoritmos que intentan solucionar el problema del apareo[ZEL99][ZEL98][ZEL00][KAN94], y algoritmos que tratan el problema del reconocimiento de rostros[BRU95][HON91][CHE91]. Se cuenta con literatura sobre las diferentes técnicas para la descripción geométrica de características faciales [YUI89][HON91].

Se evaluaron técnicas basadas en la comparación de diferentes regiones de las caras, tales como ojos, nariz y boca, usando patrones de búsqueda [BRU93] que representan estas mismas regiones.

Con la combinación de estas técnicas se determinan con gran precisión los puntos homólogos del par estereoscópico.

Los resultados obtenidos utilizando diferentes algoritmos o sus combinaciones permiten la selección de un proceso adecuado para la obtención de modelos numéricos 3D óptimos. Estos modelos numéricos correspondientes a las caras permitirán emprender la tarea de reconocimiento de aquellos que representen al mismo sujeto por comparación de un modelo calculado dado, con los modelos existentes almacenados en la Base de Datos.

En las siguientes secciones se describen antecedentes en reconocimiento de rostros, mas adelante, el detalle de soluciones propuestas, y por último resultados y conclusiones.

## **2.-Antecedentes en Reconocimiento de Rostros**

A continuación describimos metodologías utilizadas en reconocimiento de caras [CHE94].

Se han desarrollado esencialmente en estos escenarios:

- ✓ Imágenes frontales estáticas.
- ✓ Imágenes de perfil.
- ✓ Series de imágenes (dinámicas) de tomas de video.

Referimos nuestro estudio al primer escenario. Este se puede dividir en tres grupos principales:

- Métodos de segmentación de caras.
- Técnicas para la extracción de características estadísticas (Ej. la transformada de Karhunen-Loeve), coeficientes de descomposición de valores singulares, características estructurales (ojos, nariz, labios), puntos de alta curvatura.
- Técnicas de reconocimiento e identificación, que utilizan las características anteriormente mencionadas. Se presentan como estadísticas, basadas en características o en redes neuronales.

## **3.-Solución Propuesta y Búsqueda de Resultados**

A los efectos de organizar este proyecto se describe cada etapa en busca del resultado de comparación de los modelos numéricos de la siguiente manera:

- a) Procesamiento de cada imagen monocular del par estero:
  - Normalización de imágenes monoculares.
  - Extracción de características.
- b) Puesta en correspondencia.
- c) Cálculo de Modelos Numéricos.

d) Comparación de Modelos Numéricos:

- Rototraslación de Modelos Numéricos.
- Ajuste referencial de Modelos Numéricos.
- Comparación de Modelos.

**a) Procesamiento de cada imagen monocular del par estero:**

Un conjunto importante de restricciones se deriva del hecho de que las caras tienen dos ojos, una nariz, y una boca, similares en todas las caras. Sin embargo, esto mismo puede hacer la tarea de clasificación de la cara más difícil pero facilitar la búsqueda de características. El proceso de normalización de las imágenes es parte de esta etapa.

**- Normalización de Imágenes Monoculares:**

La extracción de características debe ser de alguna forma normalizado de manera de independizarse de la posición, escala, y rotación de la cara en el plano de la imagen de la cara.

Los ojos pueden ser encontrados a través de los siguientes pasos:

1. Localizar el eje simétrico aproximado de la cara.
2. Localizar el ojo izquierdo y derecho utilizando patrones correspondientes a los ojos, en los cuales la ubicación de la pupila es conocida.
3. Ajustando el origen del sistema de coordenadas al punto medio del segmento interocular, la distancia interocular y la dirección del eje ojo-a-ojo a valores predefinidos, se alcanza una invarianza de traslación, escala y rotación,

Esta técnica está basada en el trabajo de [BAR81] sobre el apareo de patrones referida al uso de coeficientes de correlación cruzada normalizada definida por

$$CN(y) = \frac{\langle ITT \rangle - \langle IT \rangle \langle T \rangle}{\sigma(IT) - \sigma(T)}$$

Donde IT es la parte de la imagen I que se compara con T (template - patrón),

$\langle \rangle$  es el operador promedio,

ITT es el producto píxel a píxel.

$\sigma$  es la desviación estándar sobre el área de comparación.

Se utiliza un patrón del ojo izquierdo y derecho, para localizar los ojos en la imagen a ser normalizada. De cada patrón se obtienen distintas imágenes del mismo variando su escala (0,7 - 0,85 - 1 - 1,15 - 1,3). La posición de los ojos se determina buscando el máximo valor absoluto de los valores de correlación calculado entre la porción de la imagen de la cara y cada uno de los patrones.

Para lograr mejores resultados de correlación en contra de los gradientes de iluminación cada imagen, se prenormaliza dividiendo cada píxel por la intensidad promedio.

Se presentan diferentes dificultades a tener en cuenta, entre estas podemos mencionar que los ojos de diferentes personas pueden ser significativamente diferentes. Estas dificultades pueden ser reducidas utilizando correlación jerárquica, [BURT88] para ello se construyen pirámides gaussianas de la imagen prenormalizada, como de los patrones. Se calcula la correlación empezando desde los niveles de más baja resolución, reduciendo progresivamente el área de computación de un nivel a otro manteniendo solamente una pequeña área.

Sea  $\alpha$  en (0,1) e  $i = 1..N$  el nivel de la pirámide donde  $L_i$  es el nivel de resolución mas bajo.

Sea  $U_i(L_j)$  el operador que produce una imagen con la resolución de nivel  $i$  de una imagen a nivel  $j$ . A cada nivel (empezando desde el nivel 2) se computa la correlación en el píxel  $x$  solamente si se satisfacen los siguientes requerimientos.

$$U_{i+1}(C_{N_i}(x)) \geq \Theta = \max_{\theta} (\theta | 1 - \alpha_i(\theta) \geq \alpha)$$

Donde  $\alpha_i$  es la distribución acumulativa (frecuencia) de los valores de correlación en el nivel  $i$ , y  $\alpha$  es la fracción de los píxeles activos (rango del cómputo de correlación) que serán proyectados al nivel superior.

Una vez que los ojos han sido detectados se ajusta la escala utilizando el radio de escala del patrón que presentó la mejor respuesta. La posición de ambos ojos se ubica utilizando la misma técnica, con un patrón correspondiente a cada ojo.

- ***Extracción de características***

En cada imagen del par estereoscópico se ubica la posición de los ojos por correlación (normalizada), con los patrones correspondientes a cada característica. Se calculan distintos valores de correlación a distintas escalas, a partir de la posición estimativa de la región de los ojos en la imagen. Se selecciona el cómputo que tenga el mayor valor absoluto de correlación.

Consiguientemente se determina la posición de las pupilas en cada imagen, que más adelante en el proceso, servirán como parámetros para eliminar variaciones de inclinación lateral entre los modelos en comparación.

Como parte de la búsqueda de otras características relevantes de la cara se estiman las zonas de la nariz y la boca por posiciones relativas, basándose en medidas antropométricas estándar.

**b) Puesta en correspondencia:**

Los pares de imágenes estereoscópicas están en un mismo referencial y sus líneas epipolares coinciden. Se utilizan algoritmos de puesta en correspondencia basados en áreas para determinar los puntos homólogos de cada par estereo [ZEL98].

**c) Cálculo de modelos numéricos:**

Se calculan los Modelos Numéricos 3D para cada cara, obteniendo las coordenadas espaciales a partir de los puntos homólogos encontrados basado en el caso geométrico simple del problema estereo.

**d) Comparación de modelos numéricos:**

- *Rototraslación de Modelos Numéricos:*

Dadas las coordenadas del centro de las pupilas (posición teórica del centro de los ojos) de cada modelo:

De **MBD**:  $p_{11}(x_1, y_1, z_1)$ ,  $p_{12}(x_2, y_2, z_2)$  de **MO**:  $p_{21}(x_3, y_3, z_3)$ ,  $p_{22}(x_4, y_4, z_4)$

Cada punto medio del segmento interocular en ambos modelos es el origen de coordenadas del sistema referencial (Fig. 1). Los puntos correspondientes a las pupilas izquierdas tanto de MO como de MBD, se hacen coincidir a través de una traslación. Después se realiza la rotación alrededor del vector unitario  $\vec{u}$ , obtenido por el producto vectorial de los vectores que contienen los segmentos interoculares. El ángulo se obtiene por producto escalar de los segmentos.

La matriz de rotación dada por la fórmula de D'Olindez Rodríguez es:

$$R^t = e^{A\phi} = I + A\text{sen}\phi + A(1 - \cos\phi)$$

donde R es una matriz de rotación 3D y

$$A = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} \quad u = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

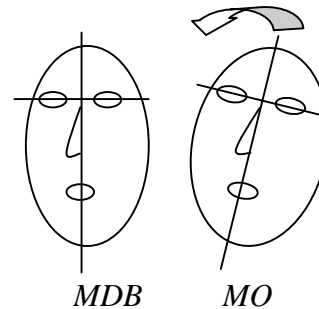


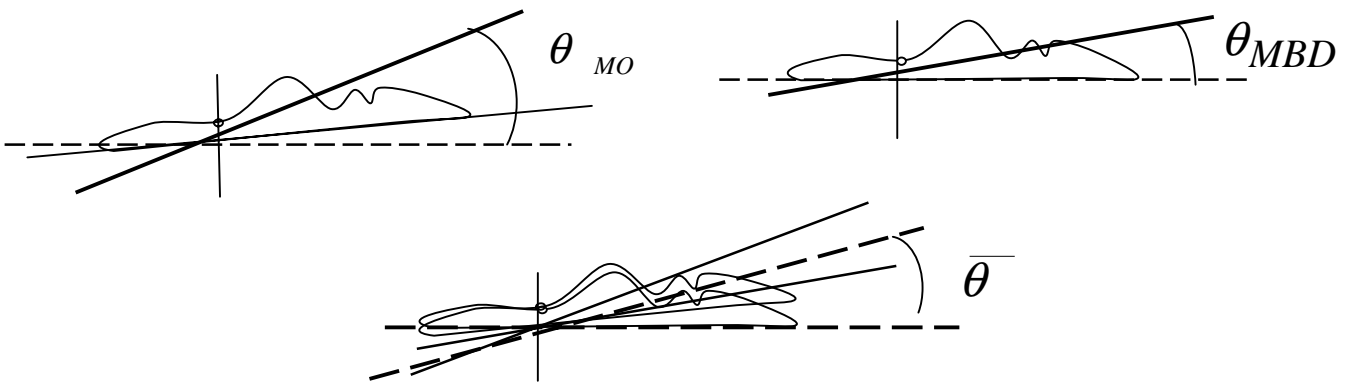
Fig. 1

El resultado será  $P' = R^T \cdot P$ , donde P y P' son los puntos del modelo numérico antes y después de la rotación.

- **Ajuste referencial de Modelos Numéricos:**

Para cada punto del modelo **MO** se obtiene el punto más cercano del modelo **MBD** en el plano  $x, y$  del referencial.

Con las coordenadas  $x, z$  de cada punto, se obtienen los ángulos que forman con respecto al origen del referencial. Para cada par de ángulos obtenidos de MO y MBD se calcula el ángulo diferencia. Una vez obtenidos todos los ángulos diferencia, se calcula el ángulo promedio.



Con este ángulo, la matriz de rotación estará dada por:  $R_y = \begin{pmatrix} \cos\theta & 0 & -\text{sen}\theta \\ 0 & 1 & 0 \\ -\text{sen}\theta & 0 & \cos\theta \end{pmatrix}$

Y por producto matricial, la rotación es:  $P_r = R_z \cdot P$

#### - Comparación de modelos:

Se procede al cálculo del error cuadrático medio entre las distancias de ambos modelos, este error determina la aprobación o rechazo de identificación, consideramos así al error como el factor de decisión.

En primera instancia, para cada punto de **MO** se busca el punto semejante correspondiente en **MBD**: Se calculan los tres puntos más cercanos en **MDB**, que formen un triángulo, con la restricción de que la recta paralela al eje z que pasa por el punto de **MO** esté contenido por dicho triángulo.

Las coordenadas del punto semejante están dadas como 'x', 'y' iguales a las de **MO** y 'z' es el promedio de los valores en 'z' de los puntos que conforman el triángulo semejante.

Entonces:  $P_{MBD} = (x, y, \frac{z_1 + z_2 + z_3}{3})$

Una vez obtenidos todos los puntos semejantes de **MBD** con respecto a **MO**, se calcula el

error cuadrático medio de las valores de 'z' :  $E = \sqrt{(z_{MBD} - z_{MO})^2}$



#### **4.- Resultados y conclusiones:**

Para la generación de Modelos Numéricos 3D de las caras se utilizaron algoritmos de puesta en correspondencia [ZEL98] basados en áreas. El uso de estos algoritmos dieron como resultado Modelos Numéricos 3D con un alto grado de precisión de cada individuo. Las imágenes estereo fueron tomadas bajo las mismas características ambientales, sin embargo, el proceso de normalización tiende a minimizar estas variaciones. Se generaron modelos numéricos 3D con varios pares de imágenes estereo que conformaron la base de datos. Cada uno de estos modelos fue comparado con un modelo numérico particular de manera de obtener distintos resultados. Cuando no se realizaban ajustes de referencial no se obtuvieron resultados satisfactorios. Después de llevar a cabo las etapas de normalización, los resultados del reconocimiento fueron satisfactorios.

El interés del método reside en la poca variación de ciertas regiones de los modelos numéricos de las caras, y la posibilidad de integrarlo con otros métodos de reconocimiento para alcanzar el modelo de respuesta perceptual de la visión humana.

Existen criterios de comparación, muchos de ellos presente en la literatura del reconocimiento de caras. En este proyecto se pretende innovar con técnicas derivadas de la Estereología, extendiendo así los métodos utilizados en el reconocimiento de caras en imágenes monoculares a modelos numéricos 3D

## .-REFERENCIAS

- [ZEL99] D.Dagun, J. Zelasco "Stereoscopic underwater images, Relative orientation and roto-rectification". Proceedings of the IASTED (SIP'99) - Oct.1999Nassau -Bahamas.
- [ZEL98] J. Zelasco, A. Donadio, D. Mendez. "Stereo Matching: Feature and area based solutions. Analysis of results". Proceedings of the IASTED – Computer Graphics and Imaging-Jun.1998-Halifax, Nova Scotia - Canada.
- [ZEL00] J.Zelasco,T. Arcomano, J.Donayo "Stereoscopic Vision, minimum set of homologous points for Rotorectification" Proceedings of the IASTED- May.2000-Pittsburg,Pensylvania–USA
- [ZEL00a] J. Zelasco, D. Dagun, J. Donayo, T. Arcomano, "Automatic Roto-Rectification of Stereo Images". Conference and Proceedings – Oceans 2000 MTS/IEEE, sept.2000.
- [BRU93] B.Brunelli and T. Poggio "Face recognition: Features Vs. Templates", IEEE Trans. Pattern Análisis and Machine Intelligence (Vol.15 – Oct. 1993)
- [BRU95] R.Brunelli and D. Falavigna "Person identificación using Multiple Cues", IEEE Trans. Pattern Análisis and Machine Intelligence (Vol.17 – Oct. 1995)
- [KAN94] T.Kanade and M.Okutomi "A Stereo Matching Algorithm With Adaptive Window : Theory An Experiment" IEEE Trans. Pattern Análisis and Machine Intelligence (Vol.16 – Oct. 1994).
- [CHE94] R. Chellappa, C. L. Wilson, S. Sirohey, "Human and Machine Recognition od Faces: A Survey". (August 1994)
- [YUI89] A. Yuille, D. Cohen, and P. Hallinan, "Feature Extraction From Faces Using Deformable Templates," in Proceedings, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 104--109, 1989.
- [HON91] Z. Hong, "Algebraic Feature Extraction of Image for Recognition," Pattern Recognition, Vol. 24, pp. 211--219, 1991.
- [CHE91] Y. Cheng, K. Liu, J. Yang, Y. Zhuang, and N. Gu, "Human Face Recognition Method Based on the Statistical Model of Small Sample Size," in SPIE Proceedings, Vol. 1607: Intelligent Robots and Computer Vision X: Algorithms and Techniques, pp. 85--95, 1991.
- [MAN92] B.S. Manjunath, R. Chellappa, and C.v.d. Malsburg, "A Feature Based Approach to Face Recognition," in Proceedings, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 373--378, 1992.
- [HAL91] P.W. Hallinan, "Recognizing Human Eyes," in SPIE Proceedings, Vol. 1570: Geometric Methods In Computer Vision, pp. 214--226, 1991.
- [KAY72] Y. Kaya and K. Kobayashi, "A Basic Study on Human Face Recognition," in Frontiers of Pattern Recognition (S. Watanabe, ed.), pp. 265--289, New York: Academic Press, 1972.
- [KIR90] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 12, 1990.
- [CHEN92] Y. Cheng, K. Liu, J. Yang, and H. Wang, "A Robust Algebraic Method for Human Face Recognition," in Proceedings, 11th International Conference on Pattern Recognition, pp. 221--224, 1992.
- [SEI91] M. Seibert and A. Waxman, "Recognizing Faces from their Parts," in SPIE Proceedings, Vol. 1611: Sensor Fusion IV: Control Paradigms and Data Structures, pp. 129--140, 1991.
- [BAR81] R. Baron, "Mechanisms of Human Facial Recognition," International Journal of Man-Machine Studies, Vol. 15, pp. 137--178, 1981.
- [BURT88] P.J. Burt "Smart sensing within a pyramid vision machine," Proc. IEEE, vol. 76, 1988.