

Reconocimiento facial robusto usando filtros de correlación diseñados a través de optimización combinatoria

Sergio Pinto-Fernández*, Alejandra Serrano-Trujillo*, Víctor H. Díaz-Ramírez* y Leonardo Trujillo Reyes**

*Instituto Politécnico Nacional – CITEDI, Ave. del Parque 1310
Mesa de Otay, Tijuana B.C., México

**Instituto Tecnológico de Tijuana, Ave. Tecnológico, Fracc. Tomás Aquino,
Tijuana B.C., México

{spinto, aserrano, vhdiaz}@citedi.mx,
leonardo.trujillo.ttl@gmail.com

Paper received on 04/10/12, Accepted on 25/10/12.

Resumen. El desempeño de los filtros compuestos de correlación para el reconocimiento de patrones depende de la adecuada selección de las imágenes de entrenamiento usadas para la síntesis de los filtros. Comúnmente, estas imágenes son elegidas de forma subjetiva por el diseñador en base su experiencia, por lo que no es posible garantizar una selección óptima. En este trabajo, se plantea el uso de un algoritmo evolutivo para la optimización combinatoria de imágenes de entrenamiento para la síntesis de filtros compuestos de correlación, usados en el reconocimiento facial. Dado un conjunto de imágenes disponibles de rostros el algoritmo encuentra la combinación óptima de imágenes de entrenamiento para la síntesis de un filtro de correlación con el mejor desempeño en términos de métricas de calidad. Como resultado, el reconocimiento facial con filtros de correlación mejora sustancialmente. Los resultados obtenidos con el algoritmo propuesto en pruebas de reconocimiento facial son presentados y discutidos en términos del desempeño con métricas objetivas y eficiencia de clasificación.

Palabras clave: Reconocimiento facial, filtros compuestos de correlación, optimización combinatoria, algoritmos evolutivos.

1. Introducción

El desarrollo de algoritmos robustos para el reconocimiento facial se ha mantenido como una línea de investigación de creciente interés, dado a la gran demanda que existe por contar con sistemas automáticos confiables para el reconocimiento de rostros. El reconocimiento facial es ampliamente utilizado en distintas aplicaciones, sin embargo, su principal área de uso se ubica en sistemas de seguridad. El problema del reconocimiento facial consiste en la validación de la identidad de una

persona, usando únicamente una imagen de su rostro. Este problema se puede abordar por distintas técnicas, por ejemplo, a través de sistemas basados en características o por sistemas basados en filtros de correlación. Los sistemas basados en características son quizás los sistemas más utilizados debido a que pueden aplicarse en una gran cantidad de problemas y comúnmente entregan buenos resultados. Sin embargo, su principal debilidad es que su desempeño general depende de la toma de decisiones subjetivas por parte del programador que podrían optimizarse mediante procedimientos formales. En los sistemas basados en filtros de correlación las coordenadas del valor de máxima intensidad en la función de salida del sistema es el estimador de máxima verosimilitud de las coordenadas del patrón objetivo en la escena observada. Comúnmente, un filtro de correlación se diseña mediante la optimización de métricas objetivas de calidad planteadas en términos de los modelos de la señal y el ruido. Un filtro de correlación compuesto, es aquel cuya respuesta al impulso está dada por una combinación de imágenes de entrenamiento que deben ser representativas del patrón objetivo y de sus diferentes versiones distorsionadas esperadas. Los filtros compuestos tienen la ventaja de que pueden construirse con el fin de obtener un desempeño optimizado para un conjunto de patrones fijos a ser identificados y para discriminar un conjunto de objetos falsos conocidos, como el fondo o cualquier otro objeto no deseado. Uno puede notar que el desempeño de un filtro de este tipo dependerá en gran medida de la correcta selección de las imágenes de entrenamiento. Aun más, esta selección puede variar considerablemente dependiendo de la aplicación. Como resultado, no es posible garantizar un óptimo desempeño al no existir un procedimiento formal para tal selección.

Existen propuestas basadas en procesos iterativos, sugeridas con el fin de resolver el problema de la selección de imágenes de entrenamiento [1,2]. Se considera que el conjunto resultante de imágenes permite la síntesis de un filtro compuesto con el mejor desempeño en términos de criterios específicos. Una limitación de las propuestas existentes es que el filtro resultante es incapaz de controlar por completo el plano de correlación a la salida, debido a que sólo se consideran imágenes del patrón objetivo para el entrenamiento. Por lo tanto, será muy probable la aparición de lóbulos laterales en el plano de correlación de salida, que son ocasionados por el fondo o cualquier otro objeto. Consecuentemente, podrían generarse errores por falsas alarmas y una baja capacidad de discriminación. Para evitar esta problemática, el diseño del filtro puede realizarse a través de un algoritmo de adaptación, utilizando en un enfoque de construcción incremental [3]. Un algoritmo de este tipo, busca patrones falsos en el área del fondo para ser rechazados en cada iteración. De esta forma, la capacidad de discriminación del filtro se incrementa monotónicamente hasta llegar a un valor aceptable. Cabe notar que el problema de la selección de las mejores imágenes de entrenamiento para un filtro compuesto es un problema de optimización combinatoria. Esto es, dadas múltiples vistas del patrón objetivo, la meta es realizar la elección de una combinación de tales vistas que, al ser utilizadas en la síntesis de un filtro compuesto, se genere el mejor desempeño con respecto a un criterio determinado, comparada con todas las otras posibles combinaciones. Por otro lado, debe observarse que la cantidad de imágenes de entrenamiento utilizadas para la síntesis de un filtro compuesto no debe incrementarse deliberadamente. Kumar y Prochavsky [4] demostraron que la relación señal

a ruido (SNR) de un filtro compuesto se reduce gradualmente al incrementar la cantidad de imágenes de entrenamiento, lo que en este planteamiento se traduce como la necesidad de obtener un conjunto pequeño de imágenes de entrenamiento que genere el mejor desempeño en términos de métricas determinadas.

En este trabajo se propone el diseño y la implementación de un algoritmo iterativo para la síntesis de filtros de correlación compuestos, optimizados para el reconocimiento facial. Dado un conjunto de imágenes de rostros, el algoritmo busca la combinación óptima de imágenes que sintetice el mejor filtro en términos de métricas de desempeño. Para ejecutar tal búsqueda, el algoritmo utiliza un enfoque evolutivo, evitando así el diseño incremental comúnmente usado.

Este documento se organiza de la siguiente manera. En la sección 2 se presenta una breve descripción de los filtros compuestos de correlación para el reconocimiento de objetos. En la sección 3, se describe la base de datos de rostros utilizada en los experimentos realizados. La sección 4 describe el algoritmo evolutivo propuesto usado para la síntesis de filtros de correlación optimizados para el reconocimiento facial. Los resultados obtenidos con el método propuesto a través de simulaciones por computadora, se presentan en la sección 5. Estos resultados, son analizados y discutidos en términos de confiabilidad en el reconocimiento facial. Finalmente, la sección 6 presenta nuestras conclusiones.

2. Filtros de correlación compuestos

En esta sección, se presenta el diseño de dos de los filtros compuestos de correlación para reconocimiento de objetos más populares: el filtro de altura máxima de correlación promedio (MACH) [5] y el filtro SDF de compromiso óptimo (OTSDF) [6,7]. Estos filtros fueron diseñados para producir picos de correlación delgados y altos en la localización del patrón objetivo en la escena observada, y para producir valores de intensidad cercanos a cero en el área ocupada por cualquier objeto desconocido. La Fig. 1 muestra el diagrama a bloques de un sistema básico de reconocimiento de objetos con filtros de correlación. El procedimiento para la detección consiste en la transformación al dominio de la frecuencia de la imagen de entrada y su espectro es multiplicado por la respuesta en frecuencia de un filtro compuesto. El resultado de esta operación se convierte al dominio espacial obteniéndose así un plano de correlación. Las coordenadas del valor de máxima intensidad, conocido como pico de correlación, indicarán la posición del objetivo.

Sea $\{T\} = \{T_i(\mu, \nu); i=1, \dots, N\}$ un conjunto de N distintas imágenes expresadas en dominio de la frecuencia, donde cada una representa una versión distorsionada del patrón a detectar $t(x, y)$. Aquí, $T(\mu, \nu)$ es la transformada de Fourier de $t(x, y)$. Los filtros compuestos deben ser capaces de reconocer al patrón objetivo $t(x, y)$ y a todas sus versiones distorsionadas en $\{T\}$ con sólo una operación de correlación.

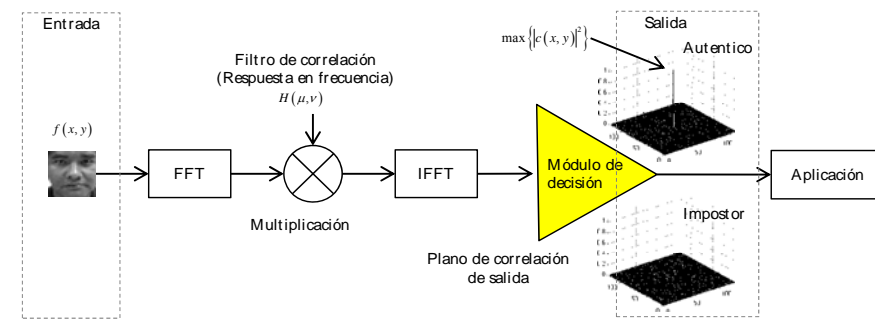


Fig. 1. Diagrama a bloques de un sistema básico de reconocimiento de objetos con filtros de correlación

2.1 Filtro de altura máxima de correlación promedio (MACH)

Un filtro MACH \mathbf{h}_{mach} se diseña para maximizar la relación entre la intensidad de la altura de correlación promedio (ACH) a la salida y la medida de similitud promedio (ASM) a través de todas las imágenes de entrenamiento [5]. En otras palabras, un filtro MACH está diseñado para maximizar la función $J = |ACH|^2 / ASM$. Sean \mathbf{X}_i y \mathbf{M} , ambas matrices diagonales de tamaño $d \times d$ que contienen los elementos de los vectores de entrenamiento \mathbf{t}_i , y el vector de entrenamiento promedio, dado por

$$\mathbf{m} = \frac{1}{N} \sum_{i=1}^N \mathbf{t}_i. \quad (1)$$

La ACH se define como el promedio de los valores de intensidad a la salida del filtro en respuesta a las imágenes de entrenamiento; esta métrica puede expresarse como

$$ACH = \frac{1}{N \cdot d} \sum_{i=1}^N \mathbf{t}_i^+ \mathbf{h}_{mach}. \quad (2)$$

Adicionalmente, la ASM puede ser vista como el error promedio entre las respuestas de correlación producidas por las imágenes de entrenamiento $\mathbf{v}_i = \mathbf{X}_i^+ \mathbf{h}_{mach}$, y la función de correlación producida por la imagen de entrenamiento promedio $\bar{\mathbf{v}} = \mathbf{M}^+ \mathbf{h}_{mach}$, es decir

$$ASM = \frac{1}{N \cdot d} \sum_{i=1}^N |\mathbf{t}_i - \mathbf{m}|^2. \quad (3)$$

Utilizando una notación matriz-vectorial, la ACH y ASM pueden ser reescritas como

$$ACH = \mathbf{m}^+ \mathbf{h}_{mach}, \quad (4)$$

y

$$\begin{aligned} ASM &= \mathbf{h}^+ \left[\frac{1}{N \cdot d} \sum_{i=1}^N (\mathbf{t}_i - \mathbf{m})(\mathbf{t}_i - \mathbf{m})^* \right] \mathbf{h}, \\ &= \mathbf{h}^+ \mathbf{S} \mathbf{h}, \end{aligned} \quad (5)$$

donde

$$\mathbf{S} = \left[\frac{1}{N \cdot d} \sum_{i=1}^N (\mathbf{t}_i - \mathbf{m})(\mathbf{t}_i - \mathbf{m})^* \right]. \quad (6)$$

El filtro MACH \mathbf{h}_{mach} se obtiene maximizando la función objetivo

$$J(\mathbf{h}_{mach}) = \frac{|ACH|^2}{ASM} = \frac{\mathbf{h}_{mach}^+ \mathbf{m} \mathbf{m} \mathbf{h}_{mach}}{\mathbf{h}_{mach}^+ \mathbf{S} \mathbf{h}_{mach}}, \quad (7)$$

donde el filtro MACH resultante está dado por [5]

$$\mathbf{h}_{mach} = \mathbf{S}^{-1} \mathbf{m}. \quad (8)$$

2.2 Filtro SDF de compromiso óptimo (OTSDF)

El filtro MACH optimiza tanto la ASM como la ACH. No obstante, en ocasiones el filtro MACH presenta problemas en la aparición de lóbulos laterales en el plano de correlación de salida, los que pueden disminuir el desempeño del filtro. Este problema puede resolverse minimizando la energía de correlación promedio (ACE), lo que provoca la generación de valores cercanos a cero en todo el plano de correlación, excepto en los valores de correlación correspondientes a la localización de las imágenes de entrenamiento. Sea \mathbf{D} una matriz diagonal de $d \times d$ elementos, en la que los valores de la diagonal principal están definidos por $E\{|\mathbf{t}_i|^2; i = 1, \dots, N\}$; es decir, por el espectro de energía promedio de las imágenes de entrenamiento en la escena. Sea \mathbf{h} cualquier filtro de correlación, por lo tanto, la ACE puede ser calculada como [6]

$$ACE = \mathbf{h}^+ \mathbf{D} \mathbf{h}. \quad (9)$$

Es importante notar que las métricas ACE y ASM están en conflicto entre sí, por lo que debe establecerse un compromiso entre ellas. El filtro OTSDF está diseñado para realizar un compromiso entre varias métricas que están en conflicto. Un filtro OTSDF que establece un compromiso entre las métricas ACE, ASM y ACH puede obtenerse minimizando la siguiente función objetivo:

$$\begin{aligned} J(\mathbf{h}_{otsdf}) &= \omega_1 ACE + \omega_2 ASM - |ACH|, \\ &= \omega_1 \mathbf{h}_{otsdf}^+ \mathbf{D} \mathbf{h}_{otsdf} + \omega_2 \mathbf{h}_{otsdf}^+ \mathbf{S} \mathbf{h}_{otsdf} - |\mathbf{h}_{otsdf}^+ \mathbf{m}|. \end{aligned} \quad (10)$$














En la Ec. (10), la ACE y ASM son funciones a ser minimizadas, ACH es una función a ser maximizada y $\omega_1 + \omega_2 = 1$ son las constantes de compromiso. El filtro OTSDF resultante, está dado por [7]

$$\mathbf{h}_{otsdf} = (\omega_2 \mathbf{S} + \omega_1 \mathbf{D})^{-1} \mathbf{m}. \quad (11)$$

3. Base de datos de los rostros

En esta sección se describe la base de datos de rostros utilizada en las pruebas realizadas con el algoritmo propuesto. Usaremos la base de datos de la CMU AMP Face Expression Database [8], que se compone de 13 diferentes sujetos. Por cada sujeto distinto existen 75 imágenes con diferentes expresiones faciales. Las imágenes son monocromáticas, de tamaño 64 x 64 píxeles y fueron tomadas bajo condiciones idénticas de iluminación, por lo que solo cambios de expresiones faciales son considerados. La Tabla 1 muestra ejemplos de diferentes imágenes de rostros por sujeto contenidas en la base de datos.

Tabla 1. Ejemplos de rostros contenidos en la CMU AMP Face Expression Database.

| | | | | | | | |
|----------|---|---|---|---|--|---|---|
| Sujeto: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Índices: | 1-75 | 76-150 | 151-225 | 226-300 | 301-375 | 376-450 | 451-525 |
| |  |  |  |  |  |  |  |
| Sujeto: | 8 | 9 | 10 | 11 | 12 | 13 | |
| Índices: | 526-600 | 601-675 | 676-750 | 751-825 | 826-900 | 901-975 | |
| |  |  |  |  |  |  | |

Sea $\{S_k\}$ una secuencia de $N_s = 75$ imágenes correspondientes al k -ésimo sujeto de la base de datos. Además, sea $\{U\} = \bigcup_{k=1}^{13} \{S_k\}$ la secuencia completa de todas las $N_U = 975$ imágenes de la base de datos. Para asegurar resultados estadísticos correctos en nuestros experimentos, hemos dividido cada uno de los conjuntos de imágenes correspondientes a cada sujeto de la base de datos en dos subconjuntos $\{S_k\} = \{A_k\} \cup \{T_k\}$ donde $\{A_k\}$ y $\{T_k\}$ son llamados el subconjunto de imágenes disponibles y el subconjunto de imágenes de prueba, respectivamente. Las imágenes disponibles $\{a_k^i(x, y) \in \{A_k\}; i=1, \dots, N_A\}$ se asumen como conocidas; por lo que todas pueden ser elegidas para el entrenamiento de los filtros. Las imágenes de prueba $\{t_k^i(x, y) \in \{T_k\}; i=1, \dots, N_T\}$ se asumen como desconocidas, por lo que sólo pueden ser usadas para probar el desempeño de los filtros. El subconjunto de imágenes disponibles se forma con $N_A = 35$ imágenes, que fueron elegidas de $\{S_k\}$ al azar, mediante el uso de una distribución uniforme. Las $N_F = N_s - N_A$ imágenes restantes se identifican como imágenes de entrenamiento. La Tabla 2 muestra la distribución de los índices de las imágenes de los rostros utilizadas en los subconjuntos $\{A_k\}$.

Tabla 2. Índice de las imágenes de la base de datos en los subconjuntos disponibles $\{A_k\}$.

| Sujeto | Índices $\{U\}$ | Índices de las imágenes $\{U\}$ utilizadas en $\{A_k\}$ |
|--------------|-----------------|--|
| $\{s_1\}$ | 1-75 | $\{A_1\}$: índices {6, 7, 8, 9, 15, 16, 17, 19, 22, 24, 26, 28, 29, 30, 33, 37, 39, 43, 47, 49, 51, 53, 54, 55, 56, 59, 60, 61, 63, 67, 69, 70, 72, 74, 75} |
| $\{s_2\}$ | 76-150 | $\{A_2\}$: índices {76, 79, 80, 81, 84, 86, 88, 89, 91, 92, 93, 95, 96, 97, 100, 104, 105, 114, 120, 121, 122, 123, 127, 131, 132, 133, 134, 137, 138, 139, 140, 141, 142, 143, 144} |
| $\{s_3\}$ | 151-225 | $\{A_3\}$: índices {152, 153, 154, 155, 157, 160, 163, 164, 165, 171, 172, 173, 174, 175, 176, 178, 179, 181, 182, 185, 186, 187, 190, 201, 204, 206, 207, 209, 211, 212, 213, 216, 219, 222, 224} |
| $\{s_4\}$ | 226-300 | $\{A_4\}$: índices {226, 228, 230, 231, 232, 234, 235, 236, 237, 238, 240, 242, 247, 249, 250, 251, 253, 261, 268, 269, 273, 274, 276, 277, 279, 281, 285, 288, 289, 290, 292, 294, 296, 298, 299} |
| $\{s_5\}$ | 301-375 | $\{A_5\}$: índices {301, 302, 305, 310, 311, 316, 317, 318, 319, 322, 323, 324, 325, 331, 333, 334, 335, 337, 340, 344, 347, 349, 350, 352, 357, 359, 363, 365, 366, 367, 369, 370, 372, 373, 375} |
| $\{s_6\}$ | 376-450 | $\{A_6\}$: índices {376, 377, 381, 382, 383, 385, 395, 401, 403, 404, 406, 407, 410, 412, 413, 414, 417, 419, 421, 422, 423, 426, 429, 430, 433, 435, 436, 439, 441, 442, 444, 445, 446, 449, 450} |
| $\{s_7\}$ | 451-525 | $\{A_7\}$: índices {452, 453, 454, 455, 459, 461, 462, 465, 466, 469, 470, 471, 473, 474, 477, 480, 482, 487, 488, 489, 493, 497, 502, 503, 504, 507, 508, 509, 510, 514, 516, 517, 519, 521, 524} |
| $\{s_8\}$ | 526-600 | $\{A_8\}$: índices {528, 529, 535, 536, 538, 539, 550, 552, 554, 556, 558, 560, 562, 563, 565, 566, 569, 573, 574, 575, 577, 580, 581, 582, 584, 585, 587, 588, 589, 593, 594, 596, 597, 598, 600} |
| $\{s_9\}$ | 601-675 | $\{A_9\}$: índices {601, 608, 609, 610, 611, 612, 613, 615, 619, 620, 624, 630, 634, 635, 638, 642, 644, 645, 646, 647, 651, 652, 654, 656, 659, 660, 661, 663, 664, 665, 666, 667, 669, 673, 675} |
| $\{s_{10}\}$ | 676-750 | $\{A_{10}\}$: índices {677, 678, 679, 680, 682, 686, 687, 688, 690, 691, 692, 694, 695, 696, 697, 701, 703, 706, 709, 713, 714, 717, 719, 720, 724, 726, 728, 732, 735, 742, 744, 746, 748, 749, 750} |
| $\{s_{11}\}$ | 751-825 | $\{A_{11}\}$: índices {751, 754, 755, 756, 757, 759, 760, 761, 763, 767, 768, 771, 772, 773, 778, 780, 781, 782, 783, 785, 786, 787, 793, 796, 797, 801, 803, 804, 806, 808, 812, 817, 821, 822, 823} |
| $\{s_{12}\}$ | 826-900 | $\{A_{12}\}$: índices {827, 829, 830, 831, 833, 834, 838, 839, 840, 845, 846, 847, 848, 854, 856, 857, 858, 859, 863, 864, 868, 869, 870, 872, 875, 881, 882, 887, 888, 889, 894, 895, 896, 898, 900} |
| $\{s_{13}\}$ | 901-975 | $\{A_{13}\}$: índices {901, 903, 904, 906, 909, 910, 911, 913, 914, 915, 917, 919, 923, 926, 927, 928, 929, 930, 933, 934, 935, 936, 937, 941, 953, 957, 958, 961, 962, 965, 967, 971, 972, 973, 975} |

4. Diseño de filtros compuestos de correlación utilizando optimización combinatoria

En esta sección se describe el algoritmo propuesto para la síntesis de filtros compuestos de correlación optimizados para el reconocimiento facial. Dado un conjunto $\{A_k\}$ de imágenes disponibles, el algoritmo busca un subconjunto óptimo $\{O_k\} \subseteq \{A_k\}$ representando las imágenes de entrenamiento elegidas. Es esperado que con las imágenes de rostros contenidas en $\{O_k\}$ se sintetice el filtro compuesto con el mejor desempeño en comparación con cualquier otra combinación de imágenes seleccionada. El desempeño del filtro puede caracterizarse en términos de una métrica de calidad específica. Para ilustrar la complejidad del problema de búsqueda, consideremos el caso de un conjunto de imágenes disponibles con $N_A = 35$ imágenes. En este caso, existen 34,359,738,367 posibles soluciones (combinaciones), por lo que una búsqueda exhaustiva sería intratable. El diagrama de bloques del algoritmo propuesto se presenta en la Fig. 2, y el procedimiento se explica de la siguiente manera:

- Paso 1: Lectura de todas las imágenes del subconjunto $\{A_k\}$.
- Paso 2: Generación de una población inicial de soluciones candidatas (combinaciones de imágenes a partir de $\{A_k\}$). Cada conjunto solución se codifica

como una cadena de longitud variable de números enteros generados al azar; de tamaño máximo N_A .

- Paso 3: Evaluar la aptitud de cada individuo de la población; esto se realiza de la siguiente manera: Sintetizar un filtro compuesto $\mathbf{h}_{current}$ utilizando las imágenes de rostros codificados en la cadena actual. A continuación, se evalúa el desempeño de $\mathbf{h}_{current}$ en el reconocimiento de rostros codificados en la cadena correspondiente. La función de aptitud busca maximizar la función

$$J(\mathbf{h}_{current}) = \frac{|ACH|^2}{\omega_1 ASM + \omega_2 ACE}. \quad (12)$$

- Paso 4: Evaluar la condición de parada, definida como un número máximo de generaciones. Si ésta condición no se ha alcanzado, continuar al paso 5.
- Paso 5: Marcar los mejores individuos dentro de la población actual, es decir, elegir los individuos con mayor aptitud (selección del 50%).
- Paso 6: Generar nuevos individuos utilizando el operador genético (cruza) basado en codificación de longitud variable [9].
- Paso 7: Reemplazar los peores individuos por los nuevos individuos generados en el paso 6.
- Paso 8: Aplicar el operador de mutación a la población (5% a cada elemento de cada individuo). Regresar al paso 3.

Puede observarse que al utilizar el algoritmo propuesto es posible encontrar un conjunto de imágenes de entrenamiento que genere el mejor desempeño en términos de un compromiso entre las métricas ACH, ACE y ASM. Esto significa que el filtro resultante tendrá un desempeño balanceado entre tolerancia a distorsiones y capacidad de discriminar patrones desconocidos. Es importante observar que para que los sujetos de la base de datos puedan ser reconocidos y clasificados a través de filtros de correlación diseñados con el algoritmo propuesto, es necesario diseñar un sistema de reconocimiento de patrones basado en un banco de filtros.

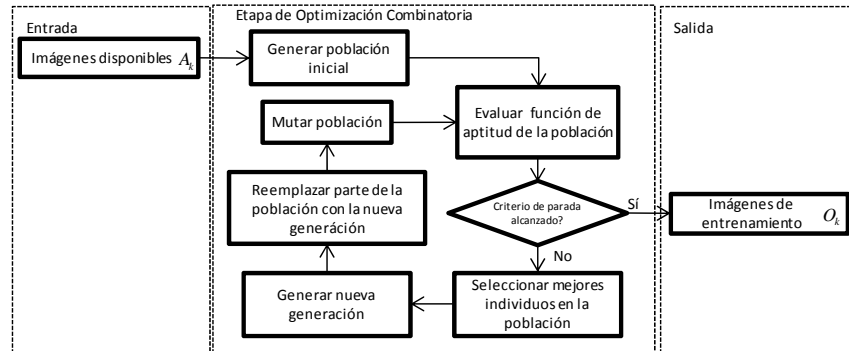


Fig. 2. Diagrama de bloques del algoritmo iterativo propuesto.

El sistema de reconocimiento debe ser capaz de validar la identidad de una imagen de entrada arbitraria perteneciente a la base de datos. El diagrama de bloques para construir un banco de filtros se presenta en la Fig. 3.

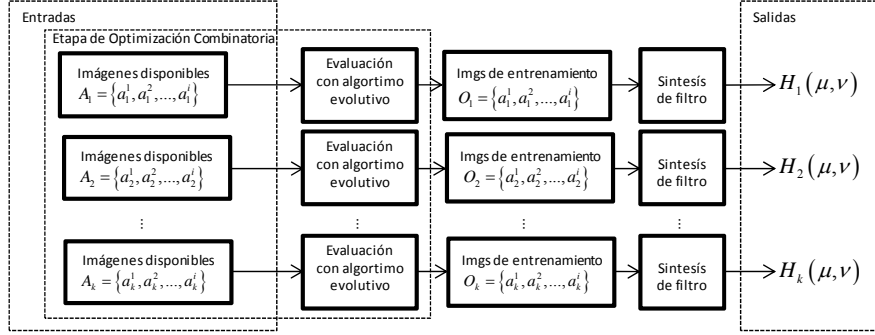


Fig. 3. Diagrama de bloques para la construcción de un banco de filtros compuestos entrenados con el algoritmo mostrado en la Fig. 2.

5. Resultados

En esta sección se presentan los resultados obtenidos con la metodología propuesta en la evaluación del reconocimiento facial con la base de datos. Los resultados obtenidos se presentan en términos del desempeño del reconocimiento facial evaluado por medio del porcentaje de eficiencia de clasificación correcta. Los resultados obtenidos con el método propuesto serán comparados con los resultados del sistema basado en características propuesto por Abusham et al [10]. Para evaluar el desempeño del algoritmo propuesto, se construyeron trece filtros de correlación (un filtro por cada sujeto) con la ayuda del algoritmo iterativo de la Fig. 2. Adicionalmente, el algoritmo se ha probado para diferentes valores de ω_1 y ω_2 . Los filtros resultantes fueron entrenados usando únicamente imágenes del conjunto $\{A_k\}$. El sistema de reconocimiento facial utilizado para evaluar el desempeño de los filtros se presenta en la Fig. 4, y el procedimiento se explica a continuación. Sea $t_k^i(x, y)$ la i -ésima imagen de prueba a lo largo del k -ésimo sujeto; se busca que esta imagen sea identificada por el sistema de reconocimiento de patrones. Primero, $t_k^i(x, y)$ es transformada al dominio de Fourier y su espectro es multiplicado por la respuesta en frecuencia de cada filtro dentro del banco $\{H_k(\mu, \nu); k=1, \dots, 13\}$. Los espectros resultantes son transformados al dominio espacial para obtener el conjunto de planos de correlación $\{cp_k(x, y); k=1, \dots, 13\}$. Posteriormente, el sistema calcula la capacidad de discriminación (DC) para cada plano de correlación de la siguiente manera:

$$DC_k = 1 - \frac{\text{MAX} \{p_i; i=1, \dots, k=13; (i \neq k)\}}{p_k}, \quad (13)$$

donde $p_k = \text{MAX} \{ |cp_k(x, y)|^2 \}$ es el valor de máxima intensidad en la función $cp_k(x, y)$. Nótese que el valor de DC para cada filtro es calculado respecto a los picos de correlación generados por los otros filtros dentro del banco. De esta manera, se espera que el filtro entrenado con las imágenes de los rostros correspondientes al sujeto de entrada mantenga un valor alto de DC. A continuación, el sistema localiza el filtro con el valor de DC más alto, y si este valor es $DC \geq 0.5$ entonces la imagen de entrada se considera reconocida, y es consecuentemente asignada a la misma categoría del sujeto para el cual fue entrenado el filtro. De otra manera, la imagen de entrada se considera rechazada.

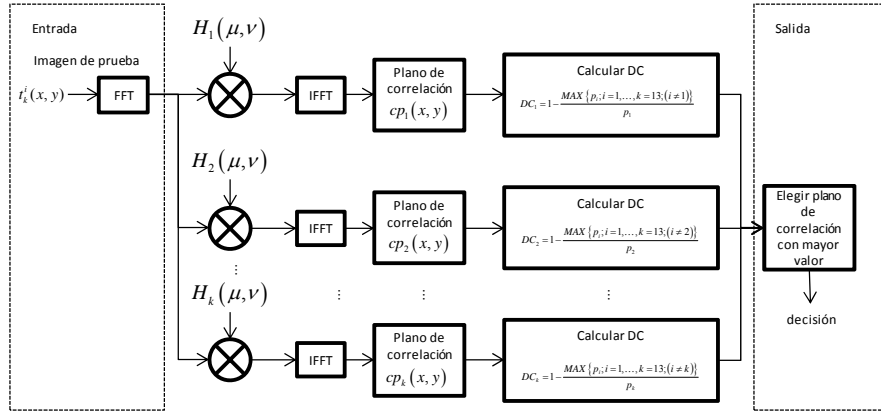


Fig. 4. Sistema de reconocimiento de patrones usado para evaluar el desempeño de los filtros compuestos de correlación entrenados con el algoritmo propuesto.

Para evaluar el desempeño del sistema propuesto, se realizaron experimentos para reconocer y clasificar las imágenes de los rostros de la base de datos. En nuestros experimentos, se utilizaron solamente las imágenes en los subconjuntos $\{T_k\}$ como las imágenes de prueba de entrada. Los resultados obtenidos se resumen en la Tabla 3, las entradas muestran el número de imágenes de entrada pertenecientes a cada sujeto (denominador) frente a las decisiones tomadas por el sistema de reconocimiento de patrones (numerador). La Tabla 3 muestra también el porcentaje de clasificación correcta obtenido en cada experimento. Los resultados obtenidos cuando $\omega_1 = 0.3$ y $\omega_1 = 0.5$ alcanzan un 100% de clasificación correcta. Para el caso de $\omega_1 = 0.7$ se ha clasificado una de las imágenes pertenecientes al sujeto 10 como si fuese el sujeto 8. Los experimentos realizados, también permiten observar el comportamiento del diseño adaptativo para los diferentes valores que puede tomar ω_1 y ω_2 . Valores de ω_1 cercanos a la unidad caracterizan al filtro con más tolerancia a las distorsiones y produce que el número de imágenes de entrenamiento necesarias, disminuyan conforme se incrementa el valor de ω_1 .

Tabla 3. Desempeño de clasificación del sistema de reconocimiento facial para valores de ω .

| | | Sujeto de entrada | | | | | | | | | | | | | | | Rechaza- dos | Correcta- mente Clasificados |
|----------|------------------------------|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------------|--------------|--------------|--------------|-------|---|-----------------|------------------------------------|
| | | | | | | | | | | | | | | | | | | |
| | | $\{T_1\}$ | $\{T_2\}$ | $\{T_3\}$ | $\{T_4\}$ | $\{T_5\}$ | $\{T_6\}$ | $\{T_7\}$ | $\{T_8\}$ | $\{T_9\}$ | $\{T_{10}\}$ | $\{T_{11}\}$ | $\{T_{12}\}$ | $\{T_{13}\}$ | | | | |
| Decisión | $\omega_1=0.3, \omega_2=0.7$ | 1 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 2 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 3 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 4 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 5 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 6 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 100 | |
| | | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 100 | |
| | | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 100 | |
| | | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 100 | |
| | $\omega_1=0.5, \omega_2=0.5$ | 1 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 2 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 3 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 4 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 5 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 6 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 1 | 0 | 0 | 0 | 0 | 100 | |
| | | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 39/40 | 0 | 0 | 0 | 0 | 97.5 | |
| | | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 100 | |
| | | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 100 | |
| | | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 100 | |
| | $\omega_1=0.7, \omega_2=0.3$ | 1 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 2 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 3 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 4 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 5 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 6 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 1 | 0 | 0 | 0 | 0 | 100 | |
| | | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 0 | 0 | 100 | |
| | | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 39/40 | 0 | 0 | 0 | 0 | 97.5 | |
| | | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 0 | 100 | |
| | | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 0 | 100 | |
| | | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 40/40 | 0 | 100 | |

En base a los resultados obtenidos, podemos observar que los filtros de correlación diseñados a través de optimización combinatoria producen altos índices de clasificación en el reconocimiento facial. Notemos que el peor índice de clasificación correcta alcanzado con el enfoque propuesto es del 97.5%. Este índice de clasificación es superior al mejor caso del sistema basado en características propuesto por Abusham et al [10], donde el índice de clasificación correcta es del 90%. En base a estos resultados, podemos decir que el reconocimiento facial con filtros de correlación es una atractiva alternativa al los sistemas basados en características.

6. Conclusiones

Se ha presentado un algoritmo iterativo basado en optimización combinatoria para la síntesis de filtros de correlación para reconocimiento facial. Se ha mostrado que dado un conjunto de imágenes de rostros, el algoritmo propuesto es capaz de encontrar una combinación óptima de imágenes de entrenamiento para la síntesis de un filtro compuesto de correlación con un desempeño optimizado para reconocimiento facial en términos de las métricas ACH, ACE y ASM. Los resultados de simulaciones por computadora obtenidos con esta propuesta exhiben altas tasas de clasificación para distintas combinaciones de parámetros del algoritmo iterativo.

7. Referencias

1. Vijaya Kumar, B.V.K., "Efficient approach to designing linear combination filters," *Appl. Opt.*, 22(10), 1445-1448 (1983).
2. Casasent, D. and Chang, W. T., "Correlation synthetic discriminant functions," *Appl. Opt.* 25(14), 2343-2350 (1986).
3. Diaz-Ramirez et al., "Multiclass pattern recognition using adaptive correlation filters with complex constraints," *Opt. Eng.* 51(3), (Marzo 2012).
4. Vijaya Kumar , B.V.K., and Pochapsky , E., "Signal-to-noise ratio consideration in modified matched spatial filters," *J. Opt. Soc. Am. A.* 3(6), 777-786 (1986).
5. Mahalanobis, A., Vijaya Kumar, B.V.K., Song, S., Sims, S.R.F., and Epperson, J.F., "Unconstrained correlation filters," *App. Opt.* 33(17), 3751–3759 (June 1994).
6. Refregier, P. "Optimal trade-off filters for noise robustness, sharpness of the correlation peak and horner efficiency". *Opt. Lett.* 16(11), 829-31 (1991).
7. Vijaya Kumar, B.V.K., Carlson, D., and Mahalanobis, A. "Optimal tradeoff synthetic discriminant function (OTSDF) filters for arbitrary devices," *Opt. Lett.* 19(19), 1556-1558 (1994).
8. "CMU AMP Face Expression Database", <<http://chenlab.ece.cornell.edu/downloads.html>>(01 Octubre 2012)
9. Inman Harvey. The saga cross: The mechanics of recombination for species with variable-length genotypes. In R. Manner and B. Manderick, (Eds.), *Parallel Problem*, pp. 269-278. North-Holland, 1992.
10. E. E. Abusham, D. Ngo, A. Teoh, "Fusion of locally linear embedding and principal component analysis for face recognition (FLLEPCA)", 3rd International Conference on Advances in Patten Recognition (ICAPR '05), vol. 3687, pp. 326–333, 2005.