

Comparación de dos Métodos de Registración Aplicados a Imágenes Retinográficas

Luciano Cárdenes¹, Amilcar Martínez¹, Adrián Salvatelli¹,
Bartolomé Drozdowicz¹

¹ Grupo de Inteligencia Artificial, Facultad de Ingeniería
Universidad Nacional de Entre Ríos
Ruta 11 Km 10 (3100), Oro Verde, Entre Ríos, Argentina
asalvatelli@bioingenieria.edu.ar

Resumen. Se implementan dos métodos de registración aplicados a pares de imágenes de fondo ocular, los cuales fueron evaluados en términos de exactitud, robustez y velocidad de ejecución. Uno de los métodos se basa en la Información Mutua, mientras que el otro en la extracción de características. Se simuló las desalineaciones entre las imágenes utilizando un modelo de transformación afín. La Información Mutua optimiza el Coeficiente de Correlación por Entropía, utilizando el método de Nelder-Mead. Se implementó además, bajo una estrategia de aceleración multiresolución. Por otro lado, el algoritmo basado en características extrae las bifurcaciones y entrecruzamientos del árbol vascular, y las corresponde mediante el coeficiente de correlación. Respecto a la exactitud, se concluye que no hay diferencias significativas entre los algoritmos. Además, la Información Mutua resultó más robusta, sin embargo el basado en características es, en promedio, 75% más rápido.

Palabras claves: registración, fondo de ojo, Información Mutua, extracción de características.

1 Introducción

La registración consiste en alinear entre sí dos o más imágenes de la misma escena. Si bien el concepto es el mismo independientemente de la aplicación, en el caso de imágenes de la retina, diversos métodos fueron adaptados y mejorados para facilitar el diagnóstico y seguimiento de patologías como la retinopatía diabética, la degeneración macular asociada a la edad y el glaucoma [1], [2]. Asimismo, la registración se utilizó para detectar cicatrices y quemaduras antes y después de la aplicación de cirugía láser [3]. Particularmente, la formación de un mosaico a partir de imágenes individuales de la retina, permite una visión más amplia de la misma facilitando el diagnóstico y seguimiento de tales patologías [4], [5], [6].

En general, los métodos de registración de imágenes pueden ser clasificados en métodos basados en el área o métodos basados en extracción de características. Los primeros optimizan una función objetivo, calculada a partir de las estadísticas de los píxeles, para alinear las imágenes. Entre estos métodos, el basado en la correlación por Transformada de Fourier, ha sido empleado para registrar imágenes de la retina

con distintos fines [7]. Otra medida de similaridad, ampliamente utilizada en la registración de imágenes, es la Información Mutua (IM) [8], que también ha sido empleada para registrar imágenes en oftalmología. En [2] se la usó para registrar 98 pares de imágenes de la retina con el fin de facilitar el estudio del nervio óptico. Por otro lado, la IM y su gradiente fueron empleados en [9] para registrar 174 pares de imágenes de un oftalmoscopio de barrido láser (SLO) con imágenes de fondo de ojo, para posteriormente fusionarlas, logrando una tasa de éxito del 94%. Alternativamente, la IM ha sido utilizada como una medida para corresponder características previamente extraídas, usando imágenes de SLO y fondos de ojo, con el fin de detectar la aparición temprana del glaucoma [10].

Por otro lado, los métodos de registración basados en la extracción de características utilizan información proporcionada por estructuras sobresalientes en las imágenes. Particularmente en los fondos de ojo, estas características están representadas por el árbol vascular (bifurcaciones), y en algunos casos el disco óptico y la mácula [11]. Sin embargo, las bifurcaciones han sido las más usadas para registrar imágenes de fondo de ojo [7], [12], [13], [14] debido a su estabilidad en el tiempo. La ventaja de éste método es realmente apreciada cuando la extracción de características es totalmente automática. No obstante, pueden no ser confiables cuando existen pocas características en el área de solapamiento entre las imágenes.

En este trabajo se implementan dos métodos de registración aplicados a pares de imágenes de fondo ocular, los cuales fueron evaluados en términos de exactitud, robustez y velocidad de ejecución. Uno de ellos se basa en estadísticas de la imagen usando la Información Mutua, mientras que el otro en la extracción de las bifurcaciones del árbol vascular, ambos usando un modelo de transformación afín.

2 Metodología

2.1 Entropía e Información Mutua

La IM mide la dependencia estadística entre dos imágenes, comparando la complejidad entre las funciones de probabilidad (fdp) conjuntas y marginales. Una alternativa a la IM muy utilizada es la Información Mutua Normalizada (IMN), que es menos sensible al área de solapamiento entre las imágenes de referencia y sensada, y a partir de la cual se determina el Coeficiente de Correlación por Entropía (ECC):

$$ECC = 2 - \frac{2}{IMN(A, B)} . \quad (1)$$

$$IMN(A, B) = \frac{H(A) + H(B)}{H(A, B)} . \quad (2)$$

Donde $H(A)$ y $H(B)$ son las entropías marginales y $H(A, B)$ es la entropía conjunta de las imágenes, ambas obtenidas de la fdp [8]. En la Fig. 1 se muestra el diagrama de flujo del algoritmo implementado.

Histograma Conjunto. El histograma conjunto $h(a,b)$ es una matriz de tamaño 256×256 , donde a y b son las intensidades de las imágenes A y B respectivamente. Los valores de gris en las imágenes varían en el rango 0 a 255. Esta matriz es inicializada con todas sus entradas iguales a cero. Luego, cada entrada en h es completada sumando un 1 (uno) en la posición $h(a,b)$, cada vez que un valor de intensidad a en la imagen A coincide con un valor b en la imagen B . Por último, el histograma es normalizado para obtener la fdp.

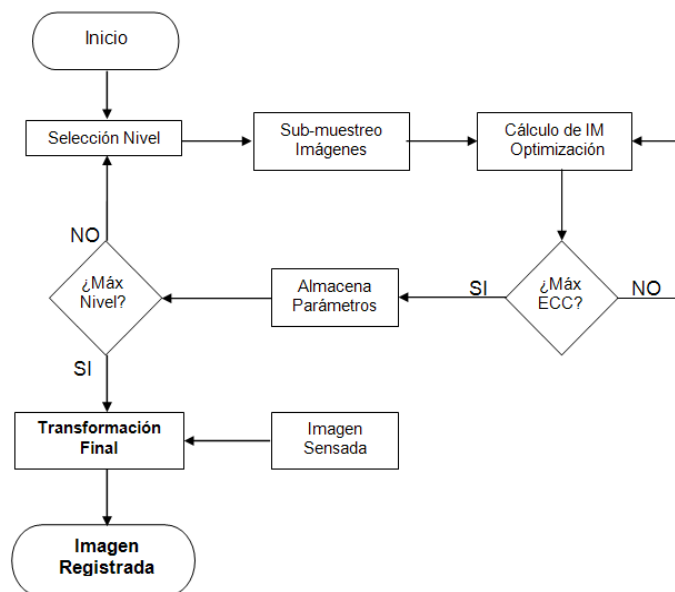


Fig. 1. Diagrama de flujo del algoritmo basado en la IM.

Método de Optimización. El método elegido en este trabajo es el Down-Hill Simplex, el cual no requiere conocer el gradiente del ECC, y permite optimizar los parámetros simultáneamente. El punto de inicialización del mismo es aquel correspondiente al 100% de factor de escala, ninguna traslación y rotación, y permite hacer la primera estimación de la transformación. A partir de éste punto se generan aleatoriamente $N+1$ puntos pertenecientes al espacio de parámetros, conformando el simplex. Luego, éste es deformado iterativamente mediante reflexiones, expansiones o contracciones de su estructura, para mover los vértices del mismo hacia el máximo de la función ECC. El punto óptimo se alcanza cuando el número de iteraciones supera las 500 o cuando la diferencia entre el mayor y el menor valor del ECC evaluada en los vértices, es igual o menor a $1e^{-4}$.

Por último, una vez recuperados los parámetros de la transformación que maximiza el ECC, la imagen sensada es transformada, concluyendo aquí el proceso de registración.

Aceleración del Método. Con el objetivo de disminuir los tiempos de cómputo del proceso de registración antes mencionado, se incorporó un esquema de registración multiresolución en dos niveles. La estrategia consiste en optimizar la función objetivo usando primero imágenes de baja resolución, obtenidas de las imágenes originales. Luego, los parámetros hallados en el nivel de baja resolución, son usados como puntos de inicio para una nueva optimización, pero seleccionando un nivel de mayor resolución. Las imágenes de baja resolución se obtuvieron submuestreándolas con factores enteros $F=2$ y $F=3$, usando interpolación de vecinos mas próximos.

2.2 Registración basada en la extracción de características

El árbol vascular es el más apropiado para extraer las características, ya que las bifurcaciones y entrecruzamientos están presentes en toda la retina y no cambian en el tiempo (salvo en algunas patologías). A continuación se explica el algoritmo basado en extracción de características.

Preprocesamiento de las Imágenes. El ajuste del contraste se realizó mediante las transformaciones morfológicas top-hat y bottom-hat. La mejora se obtiene sumando a la imagen original la imagen después de aplicarle top-hat, y al resultado se le resta la imagen obtenida con bottom-hat [15]. La transformación top-hat permite mejorar detalles en presencia de oscurecimiento, mientras que la bottom-hat remueve los detalles oscuros de la imagen manteniendo las características de brillo. El siguiente paso es segmentar el árbol vascular de acuerdo a:

$$I_{mejorada} = I_{original} - (I_{original} \circ G) . \quad (3)$$

Donde G es un elemento estructurante y “ \circ ” denota la operación de apertura. El elemento estructurante es un disco con radio 10 píxeles. La elección del tamaño guarda relación con el máximo ancho de los vasos sanguíneos. En [11] se propuso un tamaño de 5 píxeles, mientras que en [12] se utilizó uno de 9 píxeles. Por último, la estructura vascular es separada, usando el umbralamiento basado en el método de Otsu [16], la cual es afinada mediante operadores morfológicos de afinamiento para lograr un píxel de ancho, obteniéndose una imagen binarizada.

Detección de las Bifurcaciones y Entrecruzamientos. La detección de las mismas se realizó mediante la convolución entre la imagen y un kernel de 3×3 con todos sus elementos iguales a uno. Al resultado, se lo multiplicó con la imagen binarizada, para por último detectar aquellas posiciones en la imagen donde había un agrupamiento de píxeles mayor a 3. Estas coordenadas corresponden a las bifurcaciones y entrecruzamientos vasculares, con las cuales se construyen los vectores de características. Este procedimiento no asegura que todos los puntos detectados correspondan verdaderamente a bifurcaciones o entrecruzamientos. Como primera medida, se refinó el conjunto de características adoptando una vecindad de 7×7 píxeles alrededor de cada característica (en la imagen binarizada), y se sumaron los píxeles de los cuatro bordes. Si la suma era igual a 3 o 4, había una bifurcación o

entrecruzamiento, caso contrario era rechazada. De esta manera se lograron eliminar falsas características que aparecían frecuentemente.

Correspondencia de las Características. Una vez que se tienen los dos vectores de características, las mismas son correspondidas para determinar los parámetros de la transformación, utilizando el coeficiente de correlación. Se estableció una vecindad de 21×21 píxeles centrada en cada característica de los vectores de referencia y sentido. Así, cada característica del vector de referencia fue correlacionada con cada una de las características del vector sentido. La característica en el vector sentido que tuvo mayor correlación, fue retenida para formar un nuevo par con la característica del vector de referencia. Finalmente, se obtiene un nuevo vector sentido, el cual es ordenado por coeficiente de correlación descendente, seleccionando 10% de las características mejor correspondidas.

Sin embargo, el procedimiento anterior no fue suficiente para conseguir el mejor conjunto de características que permita recuperar la transformación, por lo que se calculó la transformación inversa T^{-1} utilizando el 10% de los puntos seleccionados, para mapear las características desde la imagen sensada hacia la imagen de referencia, obteniendo un nuevo vector de características. Este último es comparado con el de referencia, usando la distancia euclidiana. Aquellos puntos con una distancia menor a 5 píxeles se consideraron consistentes, mientras que el resto fue descartado. Así, se obtuvieron mejores vectores para determinar la transformación definitiva que alineará la imagen sensada con la de referencia. Finalmente, el cálculo de la transformación en sí es un procedimiento directo a partir de los vectores hallados.

3 Resultados

Se utilizaron 51 imágenes con filtro de verde y angiografías de fluoresceína con resolución de 640×480 . Los algoritmos se ensayaron simulando la desalineación entre las imágenes a registrar. De ésta manera, es posible determinar la exactitud de los mismos, ya que la transformación es conocida a priori, a diferencia de cuando se emplean imágenes desalineadas en el proceso de adquisición, para las cuales no existe un patrón de comparación. El rango de transformaciones simulado fue: -150 a 150 píxeles de traslación horizontal, -50 a 50 píxeles de traslación vertical, 0.97 a 1.03 de factor de escala y -1 a 1 grados de rotación, elegidos aleatoriamente con distribución uniforme. Por otro lado, se analiza la robustez mediante la adición de ruido, y se comparan los tiempos de ejecución de los mismos.

3.1 Exactitud

Se muestran los resultados del algoritmo basado en la Información Mutua a resolución completa (IMRC) y con la estrategia multiresolución (IMM), y el algoritmo basado en extracción de características (EC). En la Tabla 1 se presenta el error cuadrático medio de cada parámetro estimado, como así también el error total expresado en píxeles-gradados.

Tabla 1. Exactitud de los algoritmos.

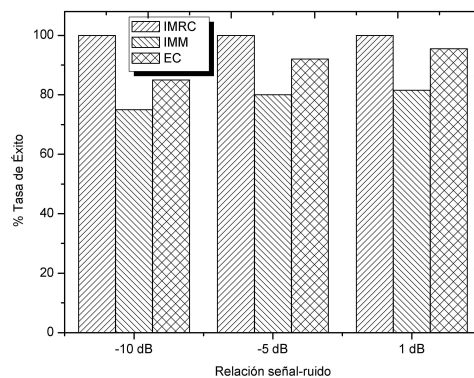
Algoritmo	e_x	e_y	e_s	e_θ	e_T
IMRC	0,0027	0,0036	0,0000	0,0005	0,0046
IMM	0,5297	0,8210	0,0007	0,0674	1,7792
EC	0,4078	0,7409	0,0400	0,0929	0,8918

En la Tabla 1 el algoritmo IMRC es el más exacto. Sin embargo, al emplear la estrategia multiresolución los errores se incrementan. Esto se debe a que en el nivel de menor resolución el algoritmo realiza una estimación gruesa de los parámetros, que luego refina en el siguiente nivel, por lo que el error es propagado. En cuanto al algoritmo EC, éste muestra errores comparables al IMM, sin embargo el error total es aproximadamente la mitad. Por otro lado, los errores en los desplazamientos horizontal y vertical son los que mas aportan al error total. Finalmente, los algoritmos brindaron buenos resultados, con una exactitud de alrededor 1.5 píxeles promedio.

3.2 Robustez

La robustez fue determinada a partir de la adición de ruido Gaussiano con media cero a las imágenes de referencia y sensada. Las simulaciones se realizaron para tres valores de SNR: -10 dB, -5 dB y 1 dB, bajo 30 transformaciones aleatorias con distribución uniforme. En el caso del algoritmo basado en IM (Fig. 2) los resultados fueron descartados siempre por una incorrecta estimación de los parámetros; mientras que en el caso del algoritmo EC, el criterio de rechazo es además, la insuficiente cantidad de características correspondidas (inferior a tres).

El algoritmo IMRC registra la totalidad de las imágenes, sin importar cuan grande sea la desalineación y el nivel de ruido existente entre las mismas, siendo este el más robusto. En cuanto al algoritmo de IMM, la tasa de éxito es la menor en todo el rango explorado de la SNR. En cambio, para el algoritmo EC, la tasa de éxito es más cercana a las del algoritmo IMRC, siendo esta similitud mayor, cuando se ensaya imágenes con filtro de verde.

**Fig. 2.** Inmunidad al ruido de los algoritmos ensayados.

3.3 Tiempo y eficiencia

El algoritmo que más tiempo demanda para registrar las imágenes es el de la IMRC, y particularmente registrando las angiografías de fluoresceína. Cuando se usa la aproximación multiresolución, los tiempos disminuyen, en promedio, un 55 % usando los dos tipos de imágenes. Este ahorro de tiempo implica una pérdida de exactitud en la estimación de los parámetros, como consecuencia de registrar las imágenes en dos niveles de resolución. Por otro lado, el algoritmo EC, es el más rápido, logrando registrar las mismas con reducciones de tiempos del 80 % con respecto al IMRC, y del 70 % con respecto a la IMM (Fig. 3).

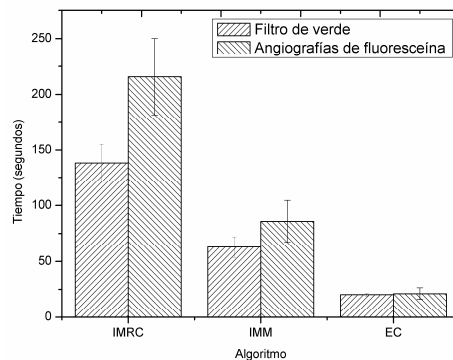


Fig. 3. Tiempos de registraci3n de los algoritmos ensayados.

4 Discusi3n

Las simulaciones permiten estudiar la capacidad de los algoritmos para recuperar los parámetros de la transformaci3n que origin3 la desalineaci3n entre las imágenes. Sin embargo, éstas no tienen en cuenta otros factores que se originan en la adquisici3n e influyen negativamente la medida de similaridad (IM) o la detecci3n de características, por ejemplo: oscurecimientos locales, regiones borrosas o muy brillantes, etc. El algoritmo IMRC recuper3 los parámetros de la transformaci3n con la mayor exactitud, registrando la totalidad de los pares, sin embargo fue el más lento de todos, en parte influenciado por la t3cnica de optimizaci3n. Si bien pueden estudiarse otros métodos de optimizaci3n, el de Nelder – Mead es una elecci3n intermedia entre robustez y velocidad de convergencia. En la Fig. 4 se muestra la performance de los algoritmos IMRC e IMM. A pesar de la p3rdida de exactitud del último, los tiempos de c3lculo se reducen considerablemente, a expensas de incrementarse el número total de evaluaciones del ECC. Por ello, en este trabajo se propuso el uso de dos niveles de resoluci3n, ya que un tercer nivel no mejor3 la velocidad de convergencia, debido a la excesiva cantidad de evaluaciones adicionales en los niveles intermedios.

El método basado en la IM puede fallar si las imágenes son de baja calidad, poco contrastadas o muy oscuras. Sin embargo es potente cuando se desea registrar regiones de dos imágenes en donde no existan características fácilmente detectables, como la mácula, la papila u otra zona con carencia de bifurcaciones vasculares. Por otro lado, y desde el punto de vista de la exactitud, el algoritmo EC es el punto intermedio entre el IMRC y el IMM. Sin embargo, el EC es un 75% más rápido registrando las imágenes. Los tiempos de registración de este último, ponen al mismo en ventaja para ser implementado en la práctica. No obstante, puede fallar si el solapamiento es muy bajo o si no existen suficientes características en común, por lo que su éxito de registración depende de la técnica de segmentación utilizada, no siendo trivial la elección de la misma, ya que el tiempo es un factor importante.

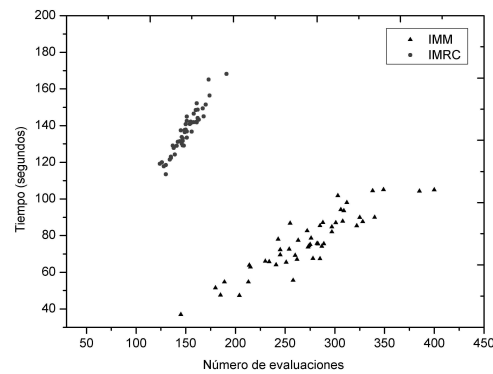


Fig. 4. Comparación entre el algoritmo IMRC y el IMM.

La Fig. 5 muestra un ejemplo de dos imágenes con un solapamiento del 50%, registrada correctamente. Trabajo a futuro: análisis de diferentes técnicas de segmentación del árbol vascular e incorporación de un modelo de transformación que tenga en cuenta la curvatura intrínseca de la retina.

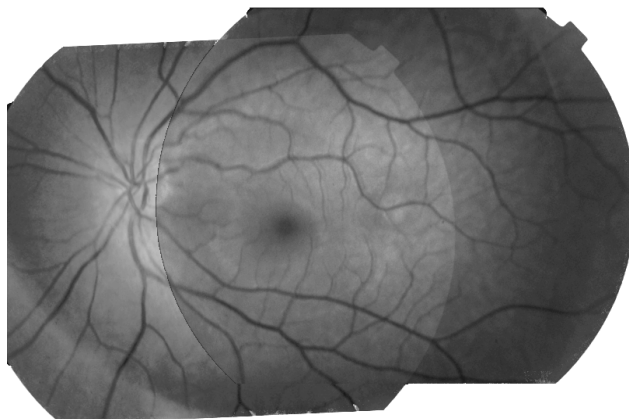


Fig. 5. Par de imágenes registrado con éxito.

5 Conclusiones

En este trabajo se han implementado dos métodos de registración aplicados a imágenes de fondo ocular: uno basado en la Información Mutua (IM) y el otro basado en la extracción de características (EC). Los algoritmos se han ensayado utilizando un total de 51 imágenes con filtro de verde y angiografías de fluoresceína, obtenidas de distintos pacientes. De las simulaciones llevadas a cabo para analizar la exactitud, se comprobó que no hay diferencias significativas entre los algoritmos siendo el error promedio 1,5 píxeles, sin embargo el IMRC es superior en exactitud comparado al de EC. En cuanto a la inmunidad al ruido, el algoritmo de IMM, tiene la menor tasa de éxito promedio 78%. El algoritmo de IMRC registró el 100% de las imágenes, mientras que el EC registró el 90%. Con respecto al tiempo de registración, el algoritmo EC es, en promedio, 75% más rápido que el basado en la IM. De los resultados obtenidos, puede inferirse que el algoritmo más apropiado para la formación de un mosaico de la retina, y más conveniente de ser llevado a la práctica es el EC, con tiempos de registración promedio de 20 segundos. De esta manera, la construcción de una única imagen de la retina que integre los distintos campos, le permitirá al médico contar con una herramienta objetiva al momento de diagnosticar y/o evaluar la evolución de patologías.

Agradecimientos. Se agradece la colaboración del Dr. Alejandro Berra, en su carácter de Asesor en la temática de análisis de imágenes de fondo ocular.

Referencias

1. Rapantzikos K., Zervakis M., Balas K.: Detection and segmentation of drusen deposits on human retina: potencial in the diagnosis of age – related macular degeneration. *Medical Image Analysis* 7, 95—108 (2003)
2. Ritter N., Owens R., Cooper J., Eikelboom R. H., Van sarloos P.P.: Registration of stereo and temporal images of the retina. *IEEE Transactions on Medical Imaging* 18, 404—418 (1999)
3. Berger J. W.: Quantitative image sequence analysis of fundus fluorescein angiography. *Journal of ophthalmic surgery and lasers* 30, 72—73 (1999)
4. Can A., Stewart C. V., Roysam B., Tenenbaum H.: A feature – based technique for joint, linear estimation of high – order image – to – mosaic transformations: mosaicing the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 412—419 (2002).
5. Becker D. E., Can A., Turner J. N., Tenenbaum H. L., Roysam B.: Image processing algorithms for retinal montage synthesis, mapping and real – time location determination. *IEEE Transactions on Biomedical Engineering* 45, 105—118 (1998)
6. Yang G., Stewart C. V., Sofka M., Tsai C.: Registration of challenging image pairs: initialization, estimation, and decision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 1973—1989 (2007)
7. Cideciyan A. V.: Registration of ocular fundus images. *IEEE Eng. Med., Biol., Mag.*, 14, 52—58 (1995)
8. Pluim J. P., Maintz J. B., Viergever M. A.: Mutual information based registration of medical images: a survey. *IEEE Transactions on Medical Imaging* 22, 986—1004 (2003)

9. Kubecka L., Jan J.: Registration of bimodal retinal images – improving modification. *IEEE Engineering in Medicine and Biology Society* 3, 1695—1698 (2004).
10. Rosin P. L., Marshall D., Morgan J. E.: Multimodal retinal imaging: new strategies for the detection of glaucoma. *Proceedings of the International Conference on Image Processing* 3, 137—140 (2002)
11. Dewan M. A., Arefin M. S., Ullah M. A., Chae O.: Automatic extraction of features from retinal fundus images. *International Conference on Information and Communication Technology ICICT 07 Dhaka, Bangladesh* 47—51 (2007)
12. Zhang E. H., Zhang Y., Zhang T. X.: Automatic image registration based on blood vessels feature point. *IEEE Proceedings of the 1st International Conference on Machine Learning and Cybernetics 2010—2015* (2004)
13. Laliberté F., Gagnon L., Sheng Y.: Registration and Fusion of retinal images – An evaluation study. *IEEE Transactions on Medical Imaging* 22, 661—673 (2003)
14. Stewart C., Tsai C. L., Roysam B.: The dual bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transactions on Medical Imaging* 22, 1379—1394 (2003)
15. Gonzalez R., Woods R., Eddins S.: *Digital image processing using matlab*. Prentice – Hall Inc. , Upper Saddle River, New Jersey (2004)
16. Otsu N.: A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics* 9, 62-66 (1979)