We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the
TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Domain-Specific Software Engineering Design for Diabetes Mellitus Study Through Gene and Retinopathy Analysis

Hua Cao, Deyin Lu and Bahram Khoobehi
Louisiana State University, University of Mississippi Medical Center, LSU Eye Center, USA

1. Introduction

Software engineering designs and practices differ widely among various application domains. This chapter is concentrating on high performance software engineering design for bioinformatics and more specifically for diabetes mellitus study through gene and retinopathy analysis. Complex gene interaction study offers an effective control of blood glucose, blood pressure and lipids. Early detection of retinopathy is effective in minimizing the risk of irreversible vision loss and other long-term consequence associated with diabetes mellitus.

Type 2 diabetes mellitus is a disorder of glucose homeostasis involving complex gene and environmental interactions that are incompletely understood. Mammalian homologs of nematode sex determination genes have recently been implicated in glucose homeostasis and type 2 diabetes mellitus. The Fem1b knockout (Fem1b-KO) mice have been developed, with targeted inactivation of Fem1b, a homolog of the nematode fem-1 sex determination gene. It shows that the Fem1b-KO mice display abnormal glucose tolerance and that this is due predominantly to defective glucose-stimulated insulin secretion. Arginine-stimulated insulin secretion is also affected. These data implicate Fem1b in pancreatic islet function and insulin secretion, strengthening evidence that a genetic pathway homologous to nematode sex determination may be involved in glucose homeostasis and suggesting novel genes and processes as potential candidates in the pathogenesis of diabetes mellitus. In addition, this chapter is going to introduce basic gene analysis approaches that can be applied on diabetes mellitus study. These approaches include searching Genbank online database using BLAST, mapping DNA, locating genes, aligning different DNA or protein sequences, determining genotypes, and comparing nucleotide or amino acid sequences using global and local alignment algorithms. Fem1b gene, as an example, is going to be discussed with these basic gene analysis approaches.

Diabetic retinopathy is the leading cause of new cases of blindness among Americans aged 20 to 64 in both predominantly white and black populations [1]. Despite the recommendation for yearly eye examinations and efforts to achieve this, of the approximately 17 million Americans with diabetes, about 6 million nationwide remain undiagnosed and untreated, or not receiving annual eye examinations, which can lead to diabetic retinopathy [2].

Early indications of retinal blood vessel abnormalities and complications provide important indicators for clinical timely diagnosis and treatment of diabetes mellitus and eye disorders. The software engineering design tool facilitates increasing the number of annual diabetic

screening eye examinations, thereby reducing the long time wait for diabetic patients to receive eye examinations. Common activities in software engineering approach for retinopathy include single or multi-modality retinal image registration, fusion, vessel pattern recognition, arteries & veins identification, and vessel diameter measurement. These methods play a major role in the development of better methods of diagnosing and treating diabetic retinopathy. Fusing the multi-modality retinal images, which usually requires intensive computational effort, is a very challenging problem because of the possible vast content change and non-uniform distributed intensities of the involved images.

This chapter is going to present a novel approach of retinal image fusion. Control points are detected at the vessel bifurcations using adaptive exploratory algorithm. Mutual-Pixel-Count (MPC) maximization based heuristic optimization adjusts the control points at the sub-pixel level. The iteration stops either when MPC reaches the maximum value, or when the maximum allowable loop count is reached. A refinement of the parameter set is obtained at the end of each loop, and finally an optimal fused image is generated at the end of the iteration. By locking the multi-modality retinal images into one single volume, the algorithm allows ophthalmologists to match the same eye over time to get a sense of disease progress and pinpoint surgical tools to increase accuracy and speed of the surgery. The new algorithm can be easily expanded to human or animals' 3D eye, brain, or body image registration and fusion.

2. Diabetes mellitus type 2 and fem1b gene

2.1 Diabetes mellitus type 2 occurrence and its diagnosis

Diabetes mellitus type 2 is a metabolic disorder that is characterized by high blood glucose in the context of insulin resistance and relative insulin deficiency [3]. The pathophysiology of Type 2 diabetes mellitus involves impaired insulin secretion, and impaired insulin action in regulating glucose and fatty acid metabolism in the liver, skeletal muscle, and adipose tissue. Many individuals with Type 2 diabetes mellitus have hypertension and perturbations of lipoprotein metabolism, as well as other manifestations of the insulin resistant syndrome. In addition to the risk for development of diabetes - specific complications of retinopathy, Type 2 diabetes mellitus is recognized as a substantial risk factor for cardiovascular disease [4].

It is recommended by the National Diabetes Data Group that diagnosis of diabetes mellitus be based on [5]

1. Two fasting plasma glucose levels of 126 mg/dL (7.0 mmol/L) or higher;
2. Two two-hour postprandial plasma glucose (2hrPPG) readings of 200 mg/dL (11.1 mmol/L) or higher after a glucose load of 75 g;
3. Two casual glucose readings of 200 mg per dL (11.1 mmol per L) or higher.

Fasting plasma glucose was selected as the primary diagnostic test because it predicts adverse outcomes (e.g., retinopathy) and is easy to perform in a clinical setting.

A mammalian Fem1 gene family, encoding homologs of fem-1, has been characterized and consists of at least three members in the mouse, designated Fem1a, Fem1b, and Fem1c; these have highly conserved homologs in humans, designated FEM1A, FEM1B, and FEM1C, respectively. Mammalian homologs of two other nematode sex determination genes, tra-2 and tra-3, have recently been implicated in glucose homeostasis and type 2 diabetes mellitus. In producing susceptibility to type 2 diabetes mellitus, NIDDM1 is known to interact with a gene, whose identity is unknown, on human chromosome 15 near the CYP19 locus at 15q21.3 [6]. This is near 15q22, where FEM1B, the human homolog of mouse Fem1b, localizes [7].

2.2 Glucose and insulin measurements

Blood glucose from tail vein was measured using an OneTouch FastTake Glucometer. Insulin was measured from plasma, tissue extracts, or cell supernatants using the Rat (Mouse) Sensitive Insulin radioimmunoassay (RIA) kit and the manufacturer's instructions. For the intraperitoneal glucose tolerance test (iPGTT), intraperitoneal insulin tolerance test (iP-ITT), and acute-phase glucosestimulated insulin secretion (A-GSIS) test, there were 12 animals in each group (12 male homozygous Fem1b-KO, 12 male wild-type controls, 12 female homozygous Fem1b-KO, and 12 female wild-type controls), aged 3 to 4 months. The arginine-stimulated insulin secretion test compared eight Fem1b-KO homozygous males with eight wild-type males, aged 6 months. D-Glucose (200 mg/ml) was administered at 2 mg/g body weight by intraperitoneal injection [8]. Tail vein blood was sampled for blood glucose determination from nonsedated animals before and at 15, 30, 60, and 120 min after glucose administration.

2.3 Fem1b-KO mice development

In this study, the gene targeting by homologous recombination has been used to generate Fem1b knockout (Fem1b-KO) mice with inactivation of the Fem1b gene. It was performed with a deletion of Fem1b coding exon 1, which contains the translation initiation codon and the first two ankyrin repeats [9]. The results show that these mice display abnormal glucose homeostasis, with abnormal glucose tolerance tests and defective glucose-stimulated insulin secretion. These findings indicate that Fem1b is involved in pancreatic islet β-cell function and provide further evidence for involvement of a pathway resembling nematode sex determination in mammalian glucose homeostasis. This approach utilized standard methodology (Figure 1) and the basic elements of the targeting vector and screening strategy by Southern blot and PCR genotyping. Figure 2 was generated by Zeiss AvioVision with the immunohistochemical analysis demonstrates a loss of specific Fem1b staining in islets of Fem1b-KO homozygotes.
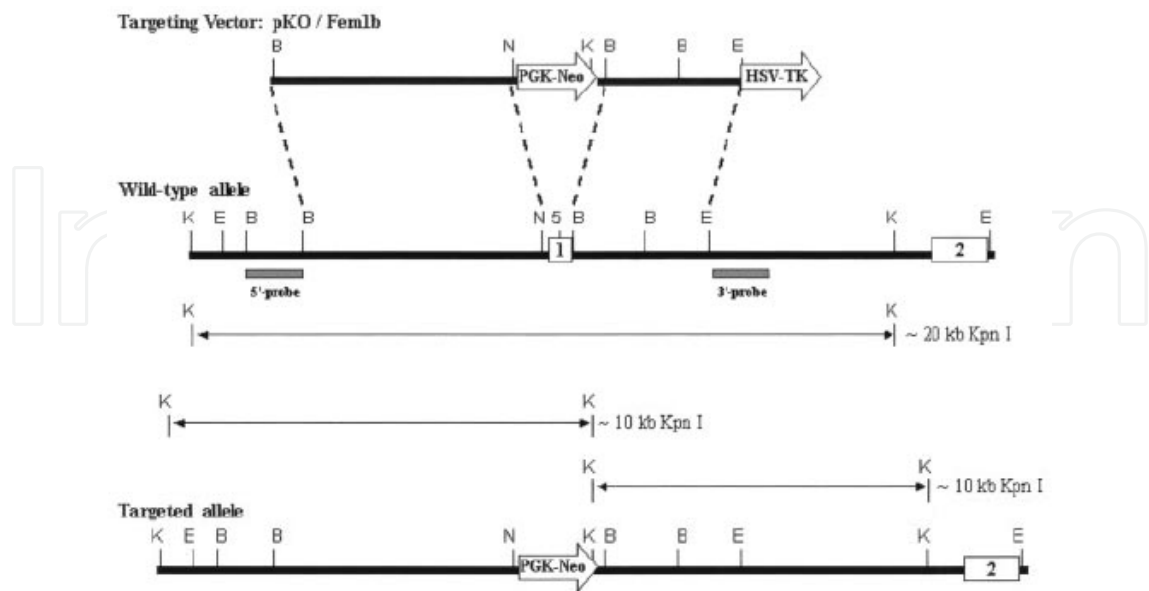

Fig. 1. General strategy, with expected KpnI digestion products along with 5_ and 3_ probes for Southern blot. The boxes labeled 1 and 2 represent exons 1 and 2, respectively. Exon 1 is replaced by the PGK-Neo gene (labeled arrow) in the targeted allele.
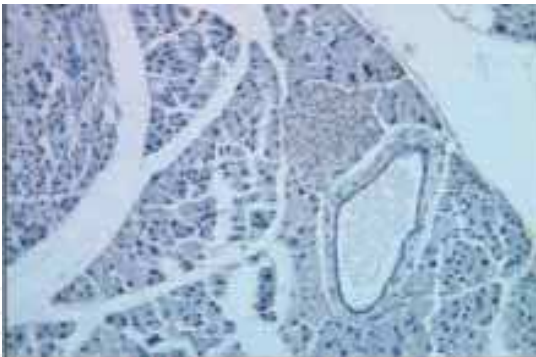
Fig. 2. Immunostaining of homozygous Fem1b-KO pancreas with anti-Fem1b C-terminus antibody Li-51, demonstrating the absence of specific staining for Fem1b.

2.4 Glucose homeostasis in Fem1b-KO mice

As noted above, mammalian homologues of nematode sex determination genes have recently been shown to be involved in glucose homeostasis and type 2 diabetes mellitus. Based on this logic, glucose homeostasis was evaluated in the Fem1b-KO mice by using established experimental methods. As a first-line screen, these mice were between 3 and 4 months of age. The iP-ITT showed minimally abnormal results (Figure 3), suggesting that insulin resistance is not the primary defect in homozygotes, although it could be contributing.
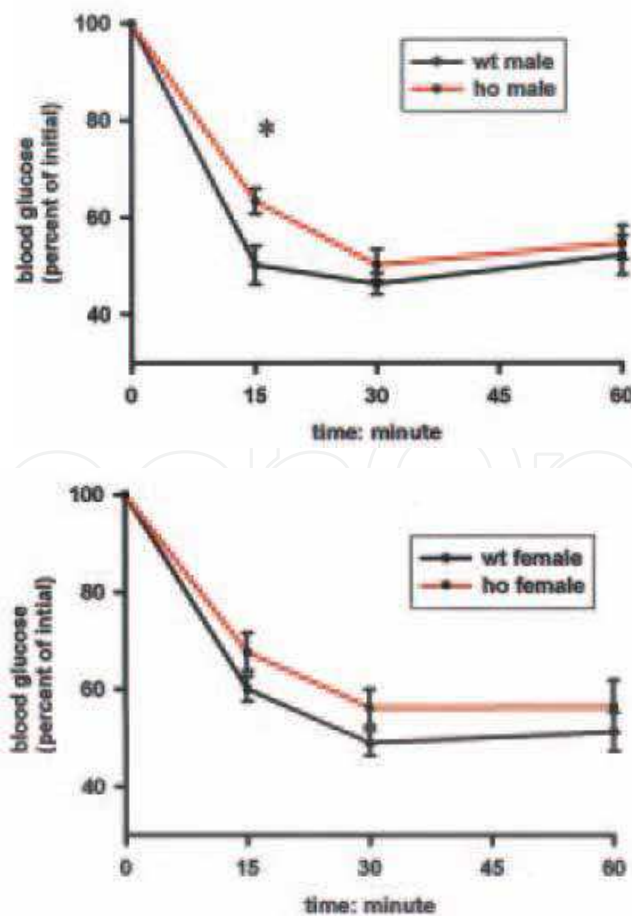

Fig. 3. Intraperitoneal insulin tolerance test (males and females).

To evaluate whether the defective acute-phase insulin secretion is related to a defect in secretion per se as opposed to a defect in insulin production, the insulin content was measured o in these mice (Figure 4), which demonstrates that Fem1b-KO homozygotes have increased insulin content compared to that of wild-type controls.
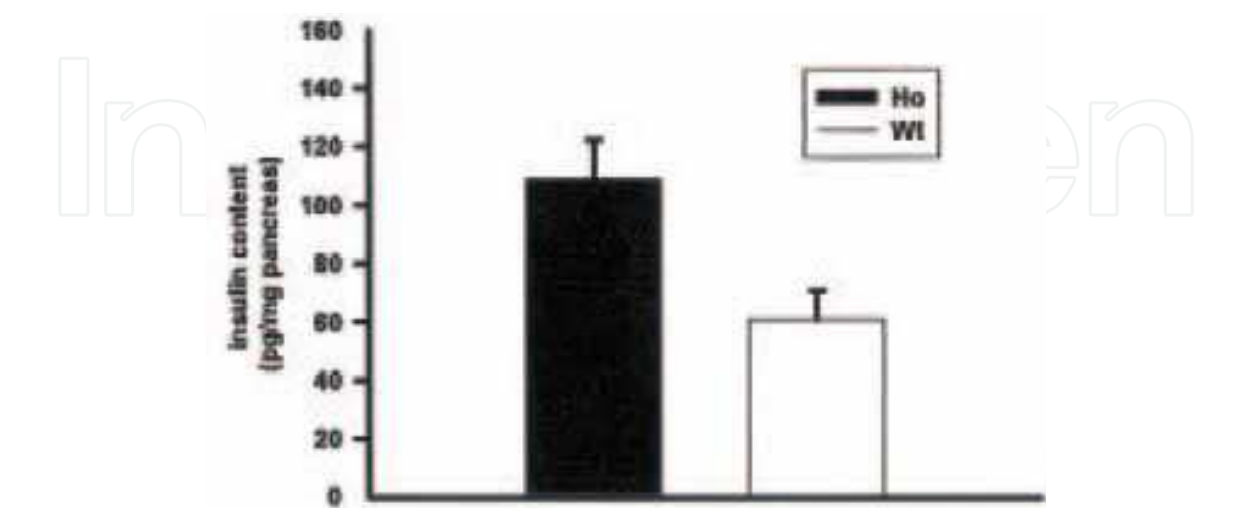

Fig. 4. Pancreatic insulin content in fasted mice (four Fem1b-KO homozygous mice and four wild-type mice).

2.5 Immunostaining of Fem1b in pancreatic islets

In humans, FEM1B has been shown to be expressed within whole pancreas [10], but cell type distribution within this organ was unknown. Immunostaining of wild-type pancreas with immunoaffinity-purified antibody shows that Fem1b protein is expressed in pancreatic islets (Figure 5).
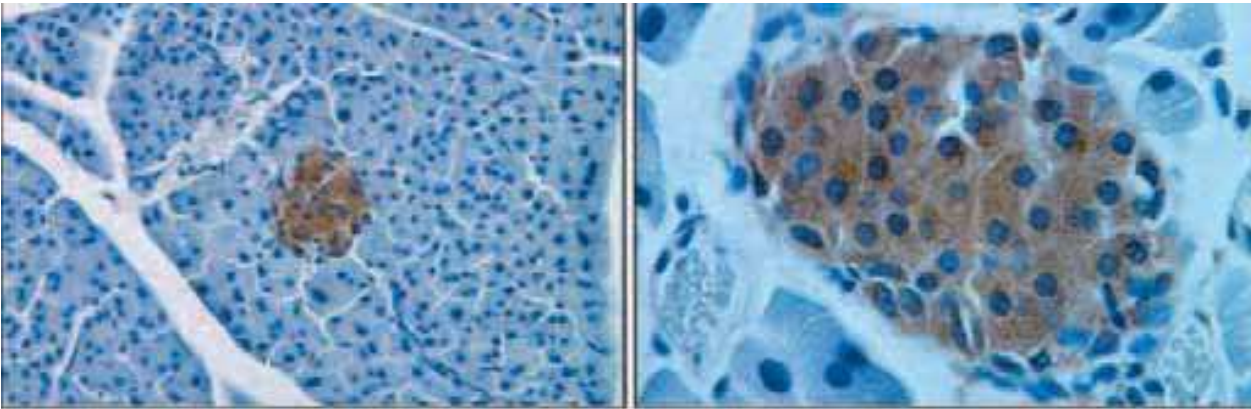

Fig. 5. Lower and higher magnification, respectively, of immunoperoxidase staining with antibody Li-51 (against the C terminus of Fem1b) counterstained with hematoxylin.

Immunostaining of the pancreas with a commercially available goat polyclonal antibody against Fem1b demonstrates the same islet staining pattern, with an absence of specific staining in the Fem1b-KO homozygotes. Coimmunostaining with antibodies to insulin, a β-cell marker, demonstrates that Fem1b is expressed in virtually all β cells (Figure 6).
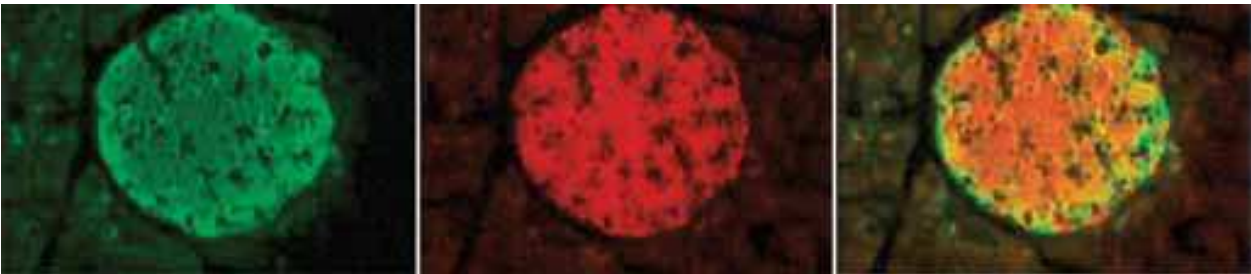
Fig. 6. Immunofluorescence staining of insulin (green), Fem1b (red), and merged image demonstrating that Fem1b is expressed not only in insulin-positive β cells but also in insulin-negative non-β cells.

The coimmunostaining with antibodies to glucagon and somatostatin, markers for α cells and δ cells, respectively, demonstrates that the Fem1b protein is also expressed in these non-β cells (Figure 7).
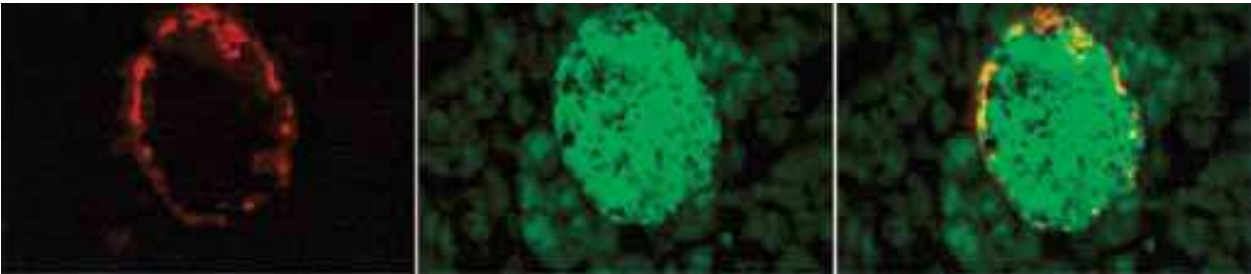

Fig. 7. Immunofluorescence staining with a combination of antibodies to glucagon and somatostatin (red) and Fem1b (green) and a merged image verifying expression of Fem1b within non- β cells.

3. Fem1b gene search and alignment

3.1 BLAST search

Basic Local Alignment Search Tool (BLAST) is an algorithm for comparing amino-acid sequences or the nucleotides. By performing a BLAST search, one is able to compare an unknown sequence with a library or database of known sequences, and identify library sequences that resemble the unknown sequence above a certain score percentage [11] (usually 40%). This chapter is going to give an example that follows the discovery of a previously unknown fem1b gene in the mouse and performs a BLAST search of the human genome to see if humans carry a similar fem1b gene. BLAST identifies sequences in the human fem1b genome that resemble the mouse fem1b gene based on similarity of sequence. Given a sequence of one fragment of mouse gene (Figure 8), BLAST software is going to search all human gene banks and find similar genes. http://blast.ncbi.nlm.nih.gov/Blast.cgi is the official BLAST website.

```
CAGAAGTCGG  ATTGAAGCCT  TGGAGCTCTT  GGGTGCCTCC  TTTGCAAATG  ATCGTGAGAA
CTATGACATC  ATGAAGACAT  ACCACTATTT  ATATTTAGCT  ATGTTGGAGA  GATTTCAGGA
TGGTGACAAC  ATTCTTGAAA  AAGAGGTTCT  CCCACCCATC  CATGCTTATG  GGAACAGAAC
TGAGTGTAGG  AACCCACAGG  AATTGGAGGC  TATTCGGCAA  GACAGAGATG  CTCTTCACAT
GGAGGGCCTT  ATAGTTCGGG  AACGGATTTT  AGGTGCTGAC  AACATTGATG  TTTCCCACCC
CATCATTTAC  AGAGGGGCTG  TCTATGCTGA  TAACATGGAG  TTCGAGCAGT  GCATCAAATT
```

Fig. 8. Unknown mouse gene.

When the results page appears, click the identifier with the highest score and you will see the following information. Here the highest score is 481. The score was calculated on the match quality and the length of the most-similar segments that occur between the unknown mouse gene and the target human fem1b gene.

```
Human - Sore: 481   Query Coverage 98%
LOCUS           NM_015322                2583 bp    mRNA    linear   PRI 26-DEC-2010
DEFINITION      Homo sapiens fem-1 homolog b (C. elegans) (FEM1B), mRNA.
ACCESSION       NM_015322
VERSION         NM_015322.3  GI:52851431
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
                Catarrhini; Hominidae; Homo.
REFERENCE       1  (bases 1 to 2583)
  AUTHORS       Ewens,K.G., Stewart,D.R., Ankener,W., Urbanek,M., McAllister,J.M.,
                Chen,C., Baig,K.M., Parker,S.C., Margulies,E.H., Legro,R.S.,
                Dunaif,A., Strauss,J.F. III and Spielman,R.S.
  TITLE         Family-based analysis of candidate genes for polycystic ovary
                syndrome
  JOURNAL       J. Clin. Endocrinol. Metab. 95 (5), 2306-2315 (2010)
   PUBMED       20200332
  REMARK        GeneRIF: Observational study of gene-disease association. (HuGE
                Navigator)
REFERENCE       2  (bases 1 to 2583)
  AUTHORS       Subauste,M.C., Sansom,O.J., Porecha,N., Raich,N., Du,L. and
                Maher,J.F.
  TITLE         Fem1b, a proapoptotic protein, mediates proteasome
                inhibitor-induced apoptosis of human colon cancer cells
  JOURNAL       Mol. Carcinog. 49 (2), 105-113 (2010)
   PUBMED       19908242
```

When you scroll down the page, you see reach a long list of the human fem1b nucleotide sequence starting with

```
ATGGAAGGACTTGCGGGGTATGTTTACAAAGCAGCGAGCGAGGGTAAAGTGCTGACCCTGG
CTGCCCTATTATTAAATCGCTCGGAATCCGATATACGATACCTTCTTGGGTACGTTAGTCA
GCAAGGAGGCCAGCGGAGTACCCCCTTGATTATAGCCGCACGTAACGGCCACGCGAAGGTC
GTGAGACTCCTACTCGAACATTATCGGGTACAAACGCAGCAAACAGGAACCGTACGGTTCG
ACGGATACGTTATAGATGGCGCGACAGCTTTATGGTGTGCCGCAGGCGCCGGTCACTTCGA
AGTAGTC
```

3.2 Sequence statistics analysis

Sections of a nucleotide sequence with a certain percentage of A+T or C+G usually indicates intergenic parts of the sequence. Figure 9 is a plot of monomer densities and combined monomer densities. One can use such statistic plot to determine if the sequence has the characteristics of a protein-coding region.

Figure 10 is the visualization of the nucleotide distribution. Figure 11 is the codon distribution showing a high amount of GAA, GAT and AAC. The amino acids for GAA, GAT and AAC are Glutamate, Aspartate, and Asparagine respectively. The corresponding bar chart distribution is displayed at figure 12. It is noticeable that it contains high volume of leucine, alanine, and valine; low volume of tryptophan, methionine, and proline.
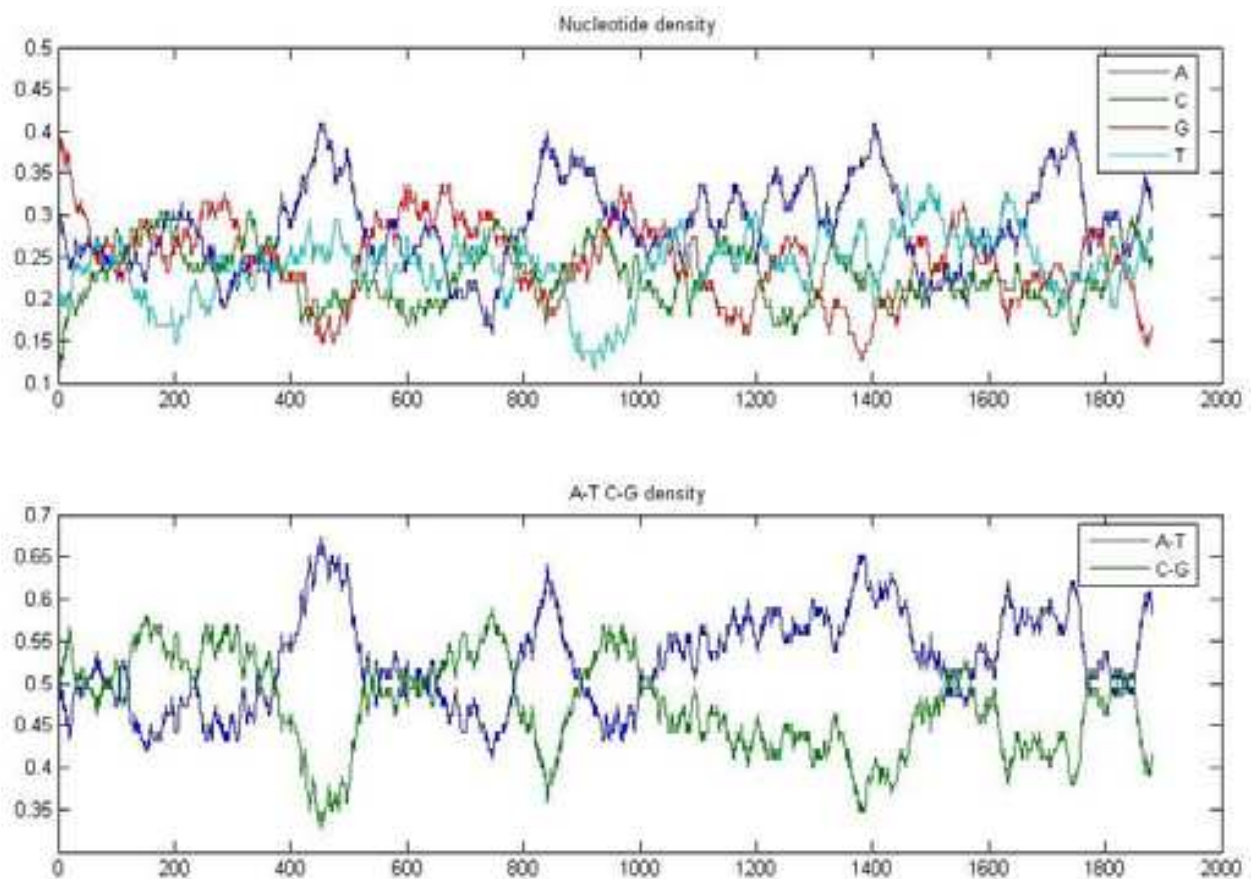
Fig. 9. Human fem1b gene's monomer densities and A-T &C-G combined monomer densities.
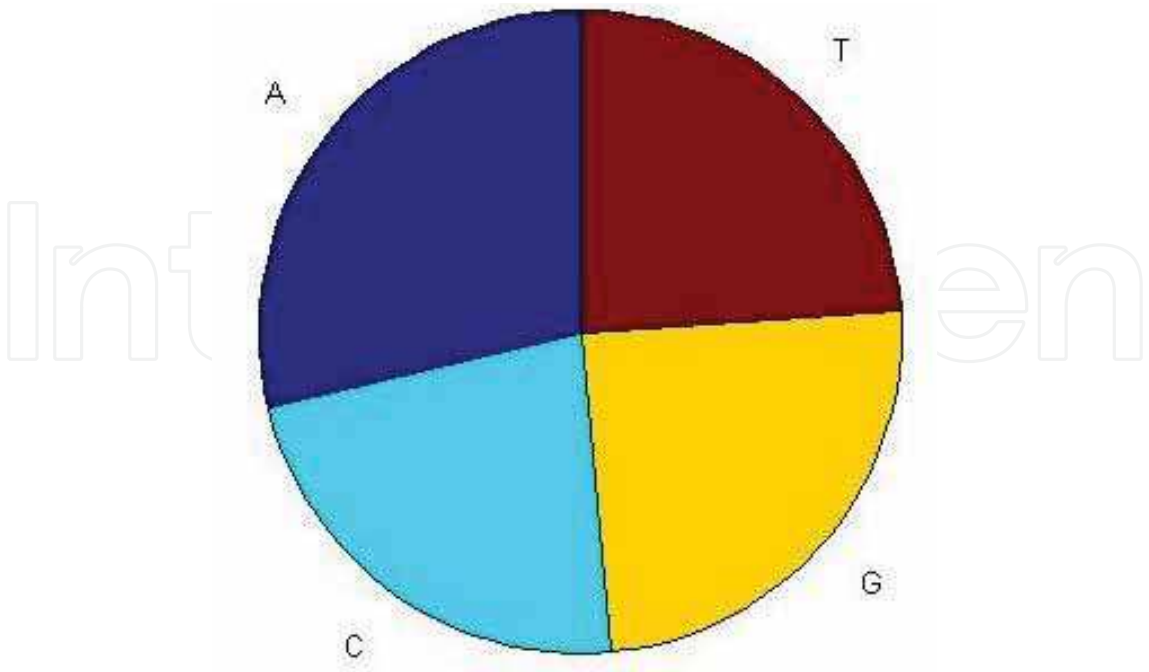

Fig. 10. Human fem1b gene's nucleotide distribution (A: 542, C: 426, G: 462, T: 451).

```
AAA  -  16        AAC  -  22        AAG  -  13        AAT  -  19
ACA  -   9        ACC  -   7        ACG  -   8        ACT  -  10
AGA  -   6        AGC  -   3        AGG  -   3        AGT  -   8
ATA  -  13        ATC  -  16        ATG  -  12        ATT  -  12
CAA  -   9        CAC  -   9        CAG  -  14        CAT  -  18
CCA  -   3        CCC  -   3        CCG  -   7        CCT  -   2
CGA  -   5        CGC  -   8        CGG  -   8        CGT  -   5
CTA  -  11        CTC  -  12        CTG  -  14        CTT  -  14
GAA  -  24        GAC  -  15        GAG  -  14        GAT  -  23
GCA  -  15        GCC  -  15        GCG  -  17        GCT  -  12
GGA  -   7        GGC  -   7        GGG  -  12        GGT  -   6
GTA  -  17        GTC  -   9        GTG  -  10        GTT  -   9
TAA  -   0        TAC  -  10        TAG  -   0        TAT  -  13
TCA  -   0        TCC  -   5        TCG  -   9        TCT  -   4
TGA  -   0        TGC  -  10        TGG  -   3        TGT  -   7
TTA  -  11        TTC  -   7        TTG  -   8        TTT  -   9
```

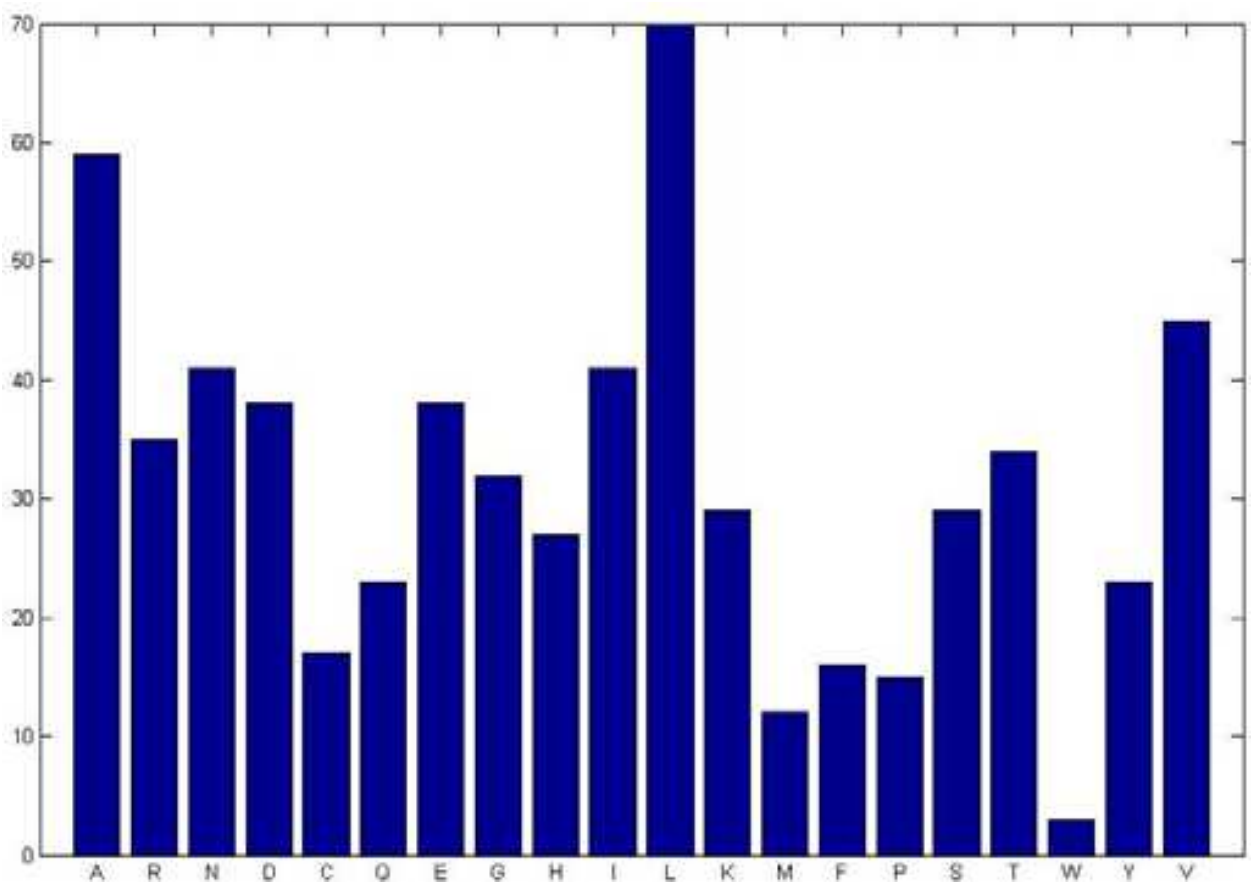Fig. 11. Human fem1b gene codon distribution.

Fig. 12. Amino acids distribution of human fem1b gene.

A: 59 R: 35 N: 41 D: 38 C: 17 Q: 23 E: 38 G: 32 H: 27 I: 41 L: 70 K: 29 M: 12 F: 16 P: 15 S: 29 T: 34 W: 3 Y: 23 V: 45

3.3 Open reading frame of Fem1b gene from both human and mouse
An open reading frame (ORF) is a nucleotide sequence without having a stop codon in a given reading frame. ORFs can be identified by examining each of the three possible reading

frames on each strand. A DNA sequence must contain a translation start codon and it is usually "AGT". Possible stop codons are "TAA", "TAG and "TGA" [11]. Identifying the start and stop codons for translation determines the ORF in a given nucleotide sequence. Once an ORF is located for a gene or mRNA, a nucleotide sequence can be translated into its corresponding amino acid sequence. Figure 13 – 15 display three reading frames for human's and mouse's fem1b gene sequences. Both genes show the longest ORF on the first reading frame.

Dot plots are one of the easiest ways to look for similarity between two sequences. The diagonal line shown in figure 16 indicates a good alignment between the human's and mouse's fem1b gene.

Fig. 13. First ORF of fem1b gene (left – mouse; right – human).

Fig. 14. Second ORF of fem1b gene (left – mouse; right – human).

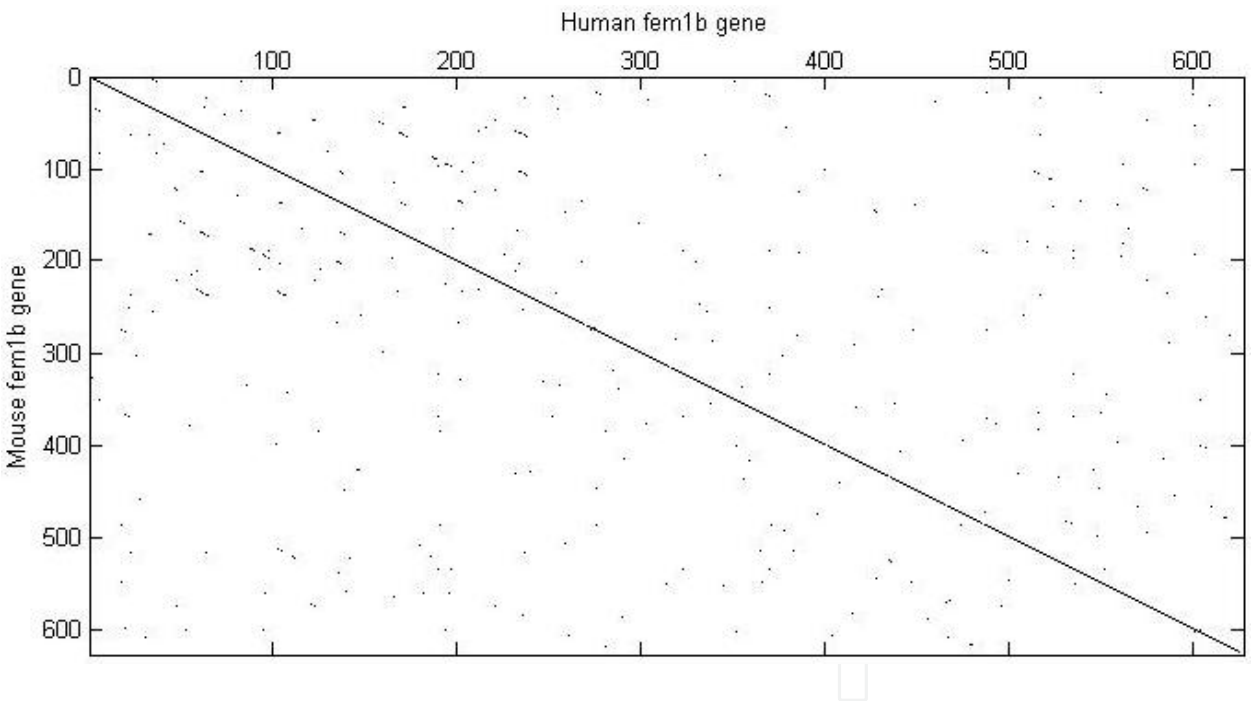Fig. 15. Third ORF of fem1b gene (left – mouse; right – human).

Fig. 16. Dot plot comparing the human and mouse amino acid sequences.

3.4 Sequence alignment
3.4.1 Global alignment
The Needleman-Wunsch algorithm, which was first published by Saul Needleman and Christian Wunsch in 1970 [12], performs a global alignment on two amino acid or nucleotide sequences. Such algorithm was the first application of dynamic programming to molecular sequence comparison. The following output was performed on two nucleotide sequences of mouse's and human's by the Needleman-Wunsch algorithm

```
Identities = 1430/1890 (76%), Positives = 1591/1890 (84%)

0001   ATGGAGGGATTGGCCGGTTACGTGTATAAGGCTGCATCCGAGGGGAAGGTGCTAACTCTAGCTG
       |||||:|||| | || || || || || ||:||:||:: |||||| ||:|||||:|| ||:||||
0001   ATGGAAGGACTTGCGGGGTATGTTTACAAAGCAGCGAGCGAGGGTAAAGTGCTGACCCTGGCTG
0065   CACTACTACTTAATCGATCGGAATCTGACATCCGCTATTTACTTGGCTATGTGTCTCAACAAGG
       | ||| || |:||||| ||||||||| || || || ||   |:||||| || || : |||:|||||
0065   CCCTATTATTAAATCGCTCGGAATCCGATATACGATACCTTCTTGGGTACGTTAGTCAGCAAGG

0129   TGGGCAAAGATCAACCCCATTGATCATCGCTGCGAGAAACGGCCATGCGAAAGTCGTTCGCCTG
       :|| ||: |:: :||||| ||||| || || ||: |:|||||||| |||||:|||||   | ||
0129   AGGCCAGCGGAGTACCCCCTTGATTATAGCCGCACGTAACGGCCACGCGAAGGTCGTGAGACTC
```

3.4.2 Local alignment

The Smith-Waterman algorithm was first published by Temple Smith and Michael Waterman in 1981 [13]. It is a well-known dynamic programming algorithm for local amino acid or nucleotide sequence alignment. Unlike the global alignment, the Smith-Waterman algorithm performs comparison among segments of all lengths and optimizes the similarity. It is guaranteed to find the optimal local alignment with respect to the scoring method. However, the Smith-Waterman algorithm requires *O(mn)* (m and n are the length of two input sequences) . In practical use, it has been replaced by the heuristic BLAST algorithm, which is much more efficient although not guaranteed to find the optimal alignments. The following output was from local alignment of the amino acid sequences of mouse's and human's using the Smith-Waterman algorithm.

```
Identities = 621/627 (99%), Positives = 627/627 (100%)

001    MEGLAGYVYKAASEGKVLTLAALLLNRSESDIRYLLGYVSQQGGQRSTPLIIAARNGHAKVVRL
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
001    MEGLAGYVYKAASEGKVLTLAALLLNRSESDIRYLLGYVSQQGGQRSTPLIIAARNGHAKVVRL

065    LLEHYRVQTQQTGTVRFDGYVIDGATALWCAAGAGHFEVVKLLVSHGANVNHTTVTNSTPLRAA
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
065    LLEHYRVQTQQTGTVRFDGYVIDGATALWCAAGAGHFEVVKLLVSHGANVNHTTVTNSTPLRAA
```

4. Diabetic retinopathy study through retinal image fusion

Hypoxia of the retina is believed to be a factor in the development of diabetic retinopathy, the leading cause of blindness worldwide. Retina image fusion provides a practical way for determination of the oxygenation status of the ocular fundus. Such method would be a valuable medical diagnostic tool for diabetic retinopathy [14], age-related macular degeneration, glaucoma [15], retinopathy of prematurity, and central retinal vein occlusion [16].

4.1 Acquisition of retinal images

Retinal images presented in this chapter were taken by a modified Topcon TRC-50EX fundus camera, with a lens and a c-mount through the vertical path of the camera. Hyperspectral images were taken through the vertical viewing port by an imaging

spectrograph and digital camera (model VNIR 100; Photon Industries Inc., Mississippi Stennis Space Center, USA) across the fundus image (Figure 17).

The subjects of the retinal images were Cynomolgus monkeys of 4 to 4.5 years of age and 2.5 to 3 kg body weight with normal eyes [17]. The use of animals for taking retinal images was approved by Louisiana State University Health Sciences Center Institutional Animal Care and Use Committee [18]. This animal usage is also conformed to the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research. The monkeys were housed in an air conditioned room with normal temperature and humidity with a 12 hour light-dark diurnal cycle.
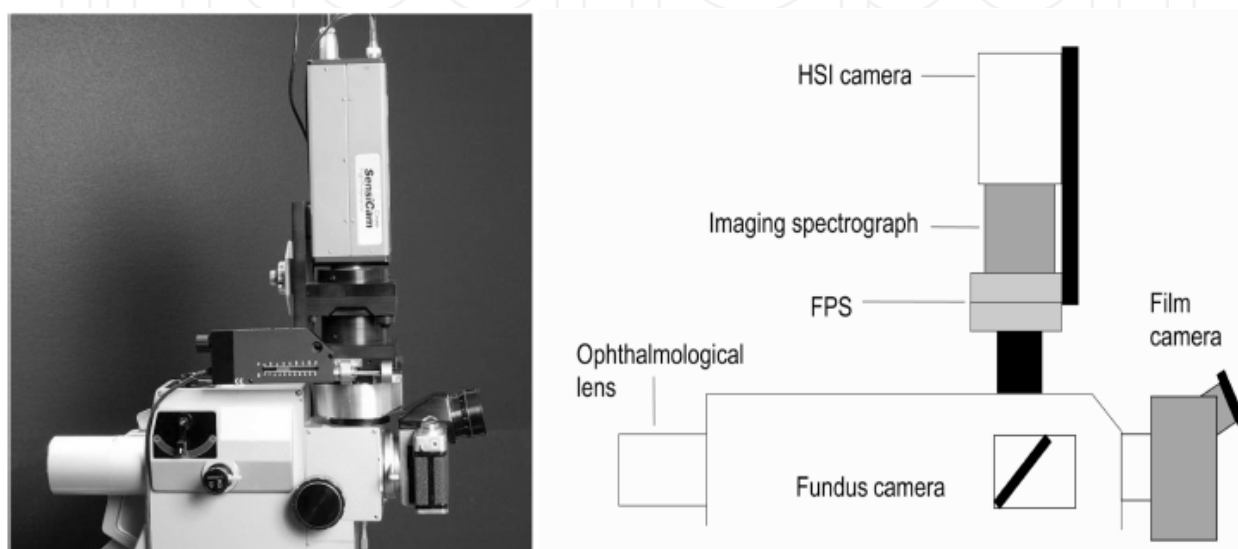

Fig. 17. Hyperspectral imaging system in relation to the fundus camera. The image is redirected upward by a mirror. The imaging system is translated over the camera port by a linear actuator mounted below the imaging spectrograph and CCD camera [17].

4.2 Retinal image fusion

There are five major steps involved in image fusion (Diagram 1):

1. The first step is the image segmentation. The segmentation subdivides an input image into its constituent regions or objects and extract/detect salient features/structure for the automated procedure.
2. The second step feature extraction is going to detect the salient structures on the target images for the feature-based approach.
3. The third step is the feature matching. The purpose of feature matching is to bringing together the information that represents same features detected at different images. The first three steps will provide the initial guess of the features for the fusion algorithm.
4. The fourth step is the optimization of the initial guess. The previously detected features will be adjusted in this step through a certain objective function.
5. The final step will transforms the images from single or different modalities into spatial alignment [19] through a certain mathematical model and then display combined view of the involved images.
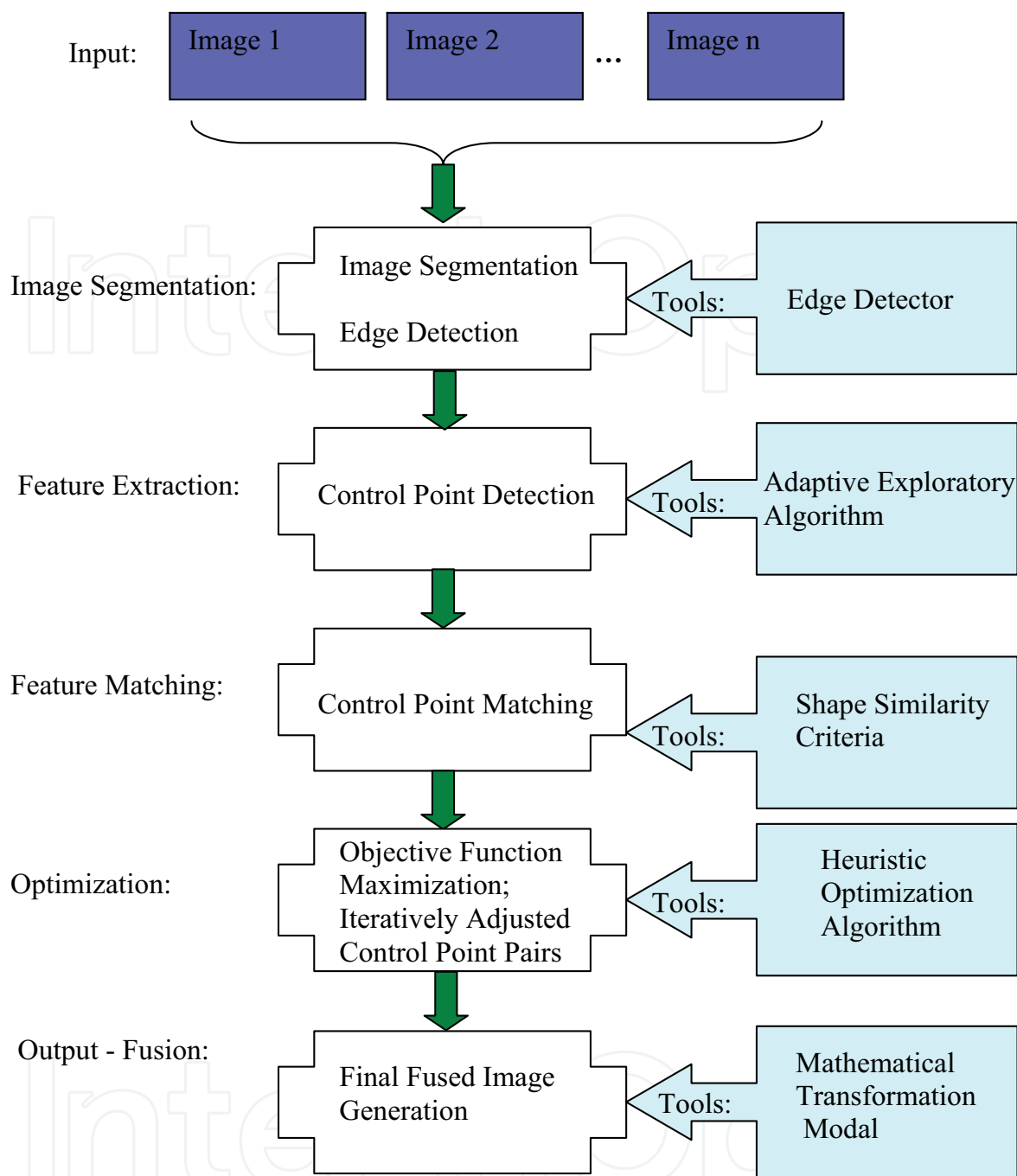
Input: Image 1   Image 2   ...   Image n

Image Segmentation: Image Segmentation
                       Edge Detection   Tools: Edge Detector

Feature Extraction: Control Point Detection   Tools: Adaptive Exploratory Algorithm

Feature Matching: Control Point Matching   Tools: Shape Similarity Criteria

Optimization: Objective Function Maximization; Iteratively Adjusted Control Point Pairs   Tools: Heuristic Optimization Algorithm

Output - Fusion: Final Fused Image Generation   Tools: Mathematical Transformation Modal

Diagram 1. Image fusion procedure.

4.2.1 Retinal vasculature extraction using canny edge detector

The Canny operator [20] is less likely than the others to be "fooled" by noise, and more likely to detect true weak edges. The Canny operator preserves most edges among all other edge detectors. Therefore, the Canny edge detector is employed in this research to extract the retinal vasculature edges. There are two criteria used in the Canny Operator to locate the rapidly changed intensity pixels. They are:

1.   Pixels where the 1st derivative of the intensity is larger in magnitude than a certain threshold;

2. Pixels where the 2nd derivative of the intensity has a zero crossing.
3. Canny's method detects edges at the zero-crossings of the second directional derivative of the image. It performs zero-crossings of

$$\frac{d^2(G \times I)}{dn^2} = \frac{d\left(\left(\frac{dG}{dn}\right) \times I\right)}{dn} \tag{1}$$

where, n is the direction of the gradient of the image; G is the edge signal; I is the image intensity. The zero-crossings of Canny's method correspond to the first directional-derivative's maxima and minima in the direction of the gradient. Edges will be identified as the maxima in magnitude. Each pixel's edge gradient is computed and compared with the gradients of its neighbors along the gradient direction. If the central pixel is smaller, mark the current edge's intensity as 0; if largest among all neighbors, keep the original intensity. Based on the nine-pixel neighborhood, the normal to the edge direction has two u_x and u_y. In order to estimate the gradient on the discrete sampling, two pixels closest to u are selected. A plane can be identified by the gradient magnitudes of three pixels. By using this plane, the gradient magnitude and the intensity at each pixel on the line can be locally estimated. The gradient magnitude at $P_{x+1, y+1}$ and $P_{x-1, y-1}$ (Figure 18) can be calculated as:

$P_{x+1, y+1}$: $$G(P_{x+1,y+1}) = \frac{u}{u_y} G(x+1, y+1) + \frac{u_y - u_x}{u_y} G(x, y+1) \tag{2}$$

$P_{x-1, y-1}$: $$G(P_{x-1,y-1}) = \frac{u}{u_y} G(x-1, y-1) + \frac{u_y - u_x}{u_y} G(x, y-1) \tag{3}$$

If the gradient at $P_{x,y}$ is greater than both of $G(P_{x+1,y+1})$ and $G(P_{x-1,y-1})$, $P_{x,y}$ will be identified as a maximum.
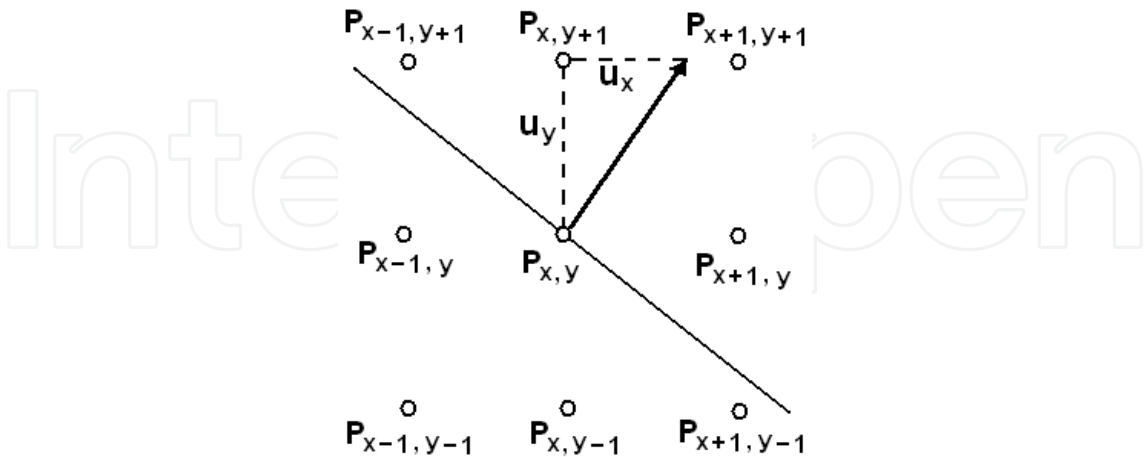

Fig. 18. Canny Edge Detection – Localization of Maxima [34].

In order to make the localization of magnitude maxima accurate, Canny defined a filter by optimizing a performance index which enhances real positive and real negative. The filter is used to minimize the probability of non-detected edge points and false detection.

$$SNR = \frac{\left| \int\limits_{-w}^{w} G(-x)f(x)dx \right|}{n_0 \sqrt{\int\limits_{-w}^{w} f^2(x)dx}} \qquad (4)$$

where, SNR stands for Signal-to-Noise Ratio; *f* is the filter; denominator is the RMSE response to noise *n(x)*. The identification of the real edge localization is defined as:

$$Localization = \frac{1}{\sqrt{E[x_0^2]}} = \frac{\left| \int\limits_{-w}^{w} G'(-x)f'(x)dx \right|}{n_0 \sqrt{\int\limits_{-w}^{w} f'^2(x)dx}} \qquad (5)$$

Two adaptive thresholds are used in Canny's method. They are high threshold and low threshold. The high threshold is used to find the start point of strong edges. Any points that meet the high threshold will be selected as the edge point. These start points are growing into different directions as long as there is no edge strength falling below the low threshold. Figure 19 is the 3D shaded surface plots of the original retinal angiogram image. The *X-Y* axis corresponds to the original image size. The height *Z* axis is a single-valued function defined over a geometrically rectangular grid. *Z* specifies the color data as well as surface height, so color is proportional to surface height with range of [0, 1] of each pixel on the image. All the retinal salient features are preserved in the Canny edges. Those salient features are the retinal vessel bifurcations, from which the control points will be selected using the Adaptive Exploratory Algorithm. Figure 20 and 21 show the retinal vessel edges detected by the Canny operator.
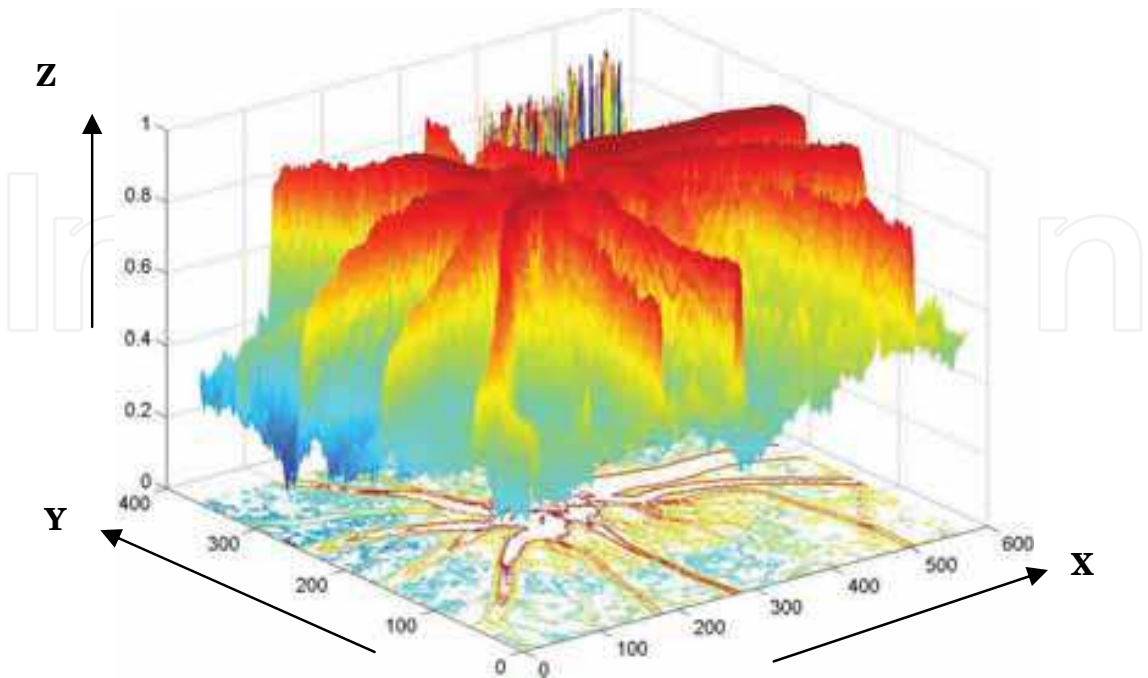

Fig. 19. 3D shaded surface plot of the angiogram image.

4.2.2 Control point detection

A good-guess of the initial control point selection ensures fused image generated at an efficient computational time. Bad control point selection will significantly increase the computation cost, or even cause the image fusion fail. Vessels or some particular abnormalities make images not necessarily matching the retina structures. Even when structure and function correspond, the abnormality still happens sometimes if inconsistence exists between structural and functional changes. Further more, angiogram images usually have higher resolution and are rich in information, whereas fundus images have lower resolution and are indeed abstract with some details or even missing some small vessels. Practically, those situations are unavoidable and will create difficulties in extracting the control points because the delineation of the vein boundaries may not be precise. In this study, control points are detected using the adaptive exploratory algorithm (Figure 20 and 21) [21].
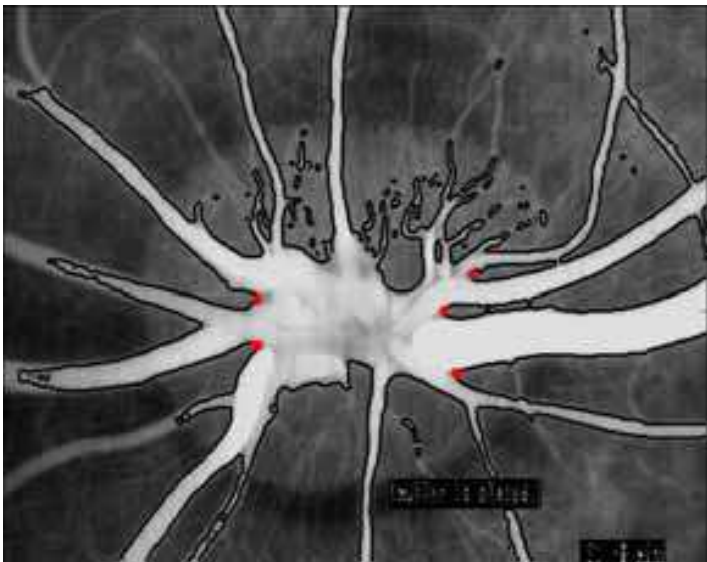
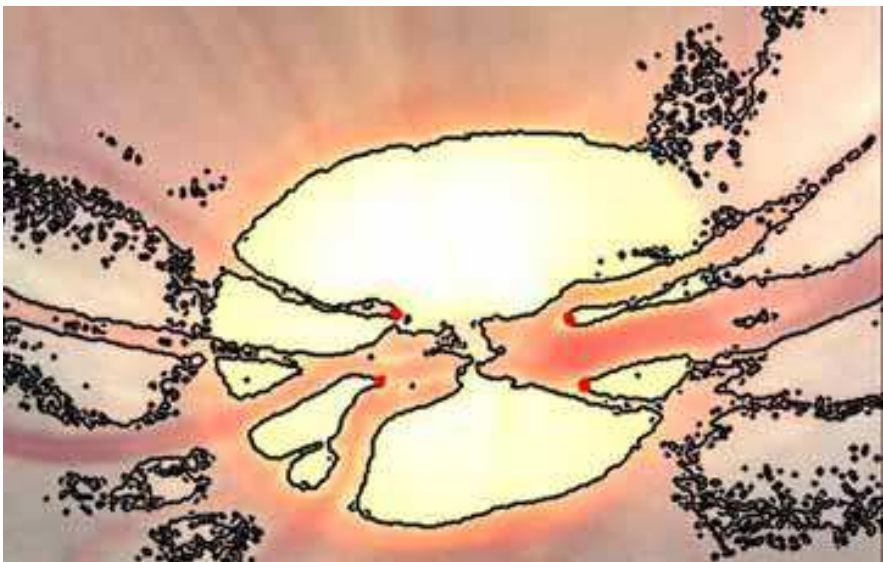Fig. 20. Angiogram grayscale reference image's control point selection.

Fig. 21. Fundus true color input image's control point selection.

4.2.3 Heuristic optimization algorithm

An optimization procedure is required to adjust the initial good-guess control points in order to achieve the optimal result. The process can be formulated as a heuristic problem of optimizing an objective function that maximizes the Mutual-Pixel-Count between the reference and input images. The algorithm finds the optimal solution by refining the transformation parameters in an ordered way. By maximizing the objective function, one image's vessels are supposed to be well overlaid onto those of the other image (Figure 23). Mutual-Pixel-Count measures the optic nerve head vasculature overlapping for corresponding pixels in both images. It is assumed that the retinal vessels are represented by 0 (black pixel) and background is represented by 1 (white pixels) in the binary 2D map. When the vasculature pixel's transformed *(u, v)* coordinates on the input image correspond to the vasculature pixel's coordinates on the reference image, the MPC is incremented by 1 (Figure 22). MPC is assumed be maximized when the image pair is perfectly geometrically aligned by the transformation. After pre-processing, the binary images of the reference and input images are obtained, i.e. I_{ref} and I_{input}. Only black pixels from both images contribute to MPC. The ideal case is that all zero pixels of the input image are mapped onto zero pixels of the reference image. The problem can be mathematically formulated as the maximization of the following objective function:

$$f_{mpc}(x,y,u,v) = \sum_{\substack{I_{ref}(x,y)=1 \ and \ I_{input}(u,v)=1}}^{u,v \in ROI} I_{input}(T_x(u,v), T_y(u,v)) \tag{6}$$

where f_{mpc} denotes the value of the Mutual-Pixel-Count. T_x and T_y are the transformations for u and v coordinates of the input image. The ROI (Region-of-Interest) is the vasculature region where the MPC is calculated on.

Coordinates' adjustment is iteratively implemented until one of the following convergence criteria is reached: (1). Predefined maximum number of loops is reached; or (2). the updated f_{MPC} is smaller than ε , i.e.

$$\left| f_{MPC^{n+1}}(x,y,u,v) - f_{MPC^n}(x,y,u,v) \right| < \varepsilon \tag{7}$$

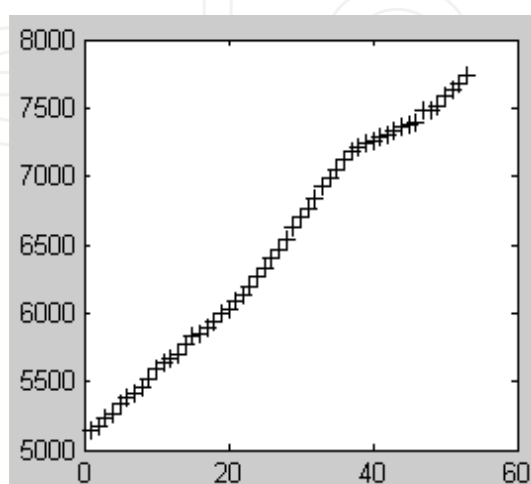where ε is a very small non-negative threshold.

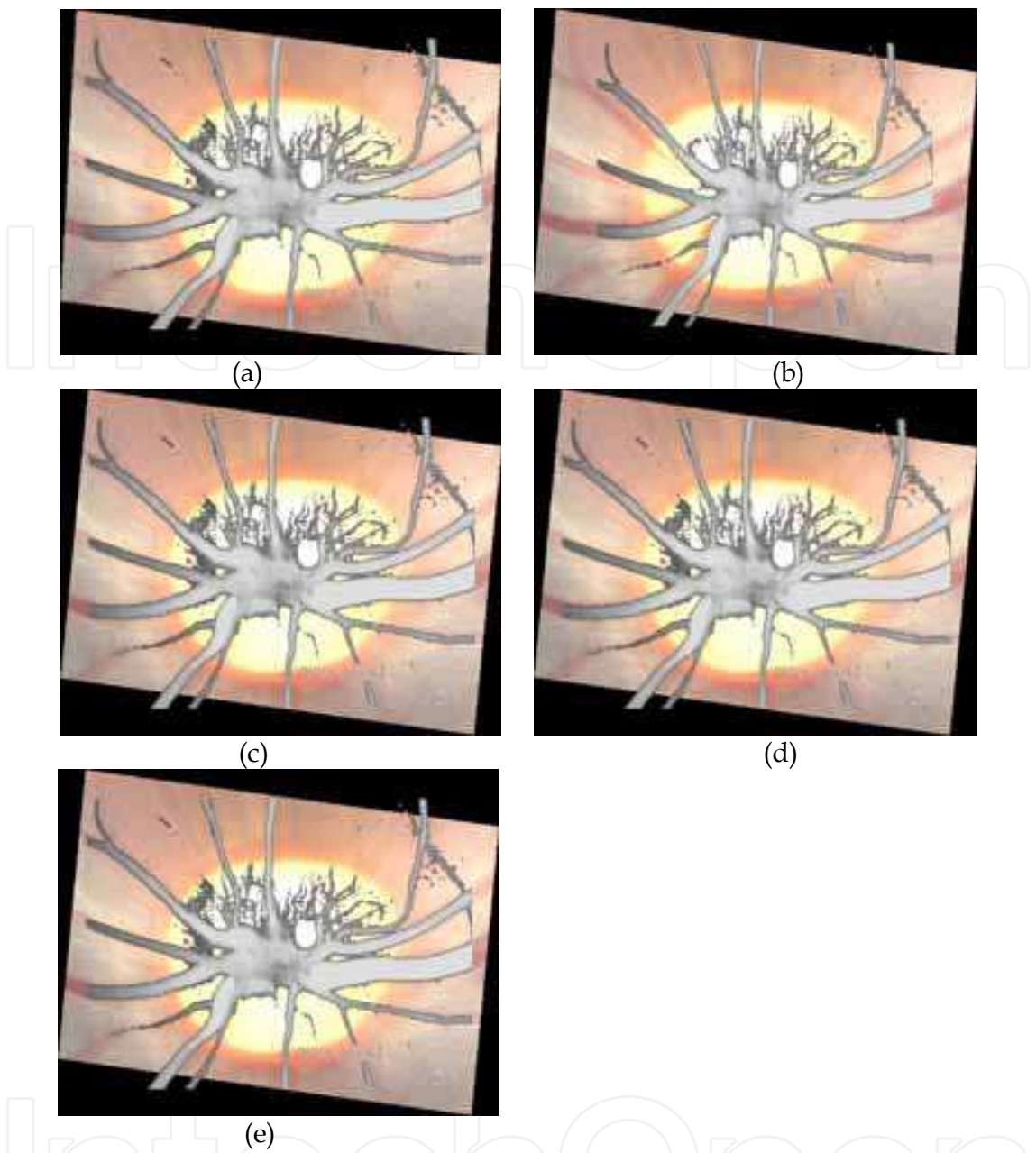Fig. 22. f_{MPC} increasing during the iteration (Y-axis is f_{MPC}; X-axis is the loop count).

(a)

(b)

(c)

(d)

(e)

Fig. 23. Fused image improvement during the iteration. From (a)-(e) fMPC = 5144, 7396, 7484, 7681, 7732.

4.2.4 Affine transformation model

A mathematical model is the tool for transforming the target images and fusing them into one single volume. Affine model is a basic geometrical transformation in image processing and is defined as Eq. 8 and 9. The DOF of the affine model is 6 because it has six parameters, i.e. $a_1, a_2, b_1, b_2, a_3,$ and a_4.

$$\begin{pmatrix} U \\ V \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & b_1 \\ a_3 & a_4 & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad U(x, y) = a_1x + a_2y + b_1 \qquad (8)$$

$$V(x, y) = a_3x + a_4y + b_2 \qquad (9)$$

Affine model's advantage lies in that it can measure lost information such as skew, translation, rotation, shearing and scaling that maps finite points to finite points and parallel lines to parallel lines (Figure 24 and Table 1). Its drawback lies in the strict requirement that at least six pairs of control points are needed [19].
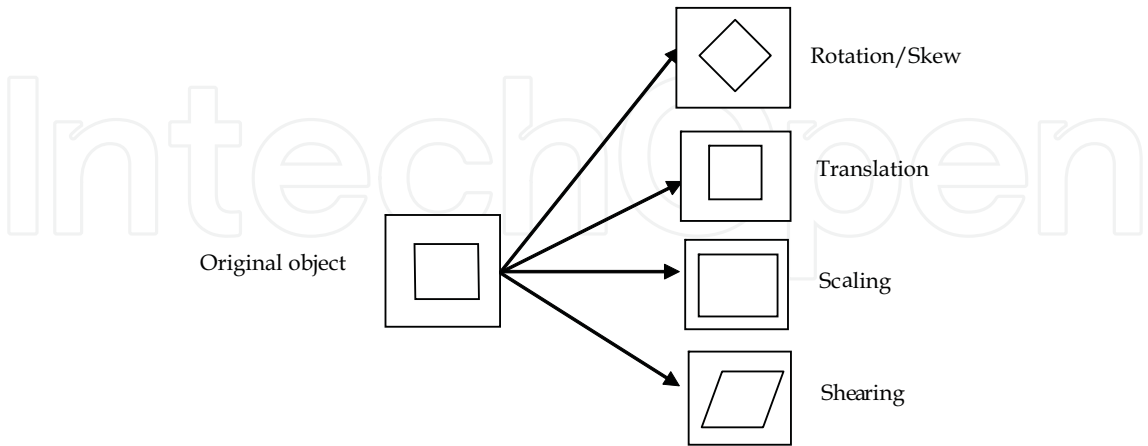

Fig. 24. Transformation Graphs.

Transformation	Description
Rotation/Skew	Points are rotated by an angle θ.
Translation	A linear shift in the position of the vertical and horizontal coordinates of the image in one plane to another set in the same spatial domain.
Scaling	A transformation of the horizontal and vertical coordinate points characterized by a certain scale factor.
Shearing	A transformation in which all points along a line remain fixed while other points are shifted parallel to the line by a certain distance proportional to their perpendicular distance from the line [22].

Table 1. Transformation Descriptions.

5. Conclusion and future directions

Fem1b functions in vivo to regulate insulin secretion and plasma glucose levels. Fem1b-KO mice do not have fasting hyperglycemia but rather have defective acutephase GSIS. Such defective acute-phase GSIS is the earliest detectable defect in humans destined to develop diabetes and may represent the primary genetic risk factor predisposing to diabetes [23]. With aging and superimposed insulin resistance, fasting hyperglycemia and overt diabetes later develop. Therefore, the Fem1b-KO mouse model is a key component of the complex pathogenesis of type 2 diabetes mellitus. Both male and female homozygotes display abnormal glucose tolerance. The role of Fem1b in pancreatic islet insulin secretion strengthens evidence that a genetic pathway homologous to nematode sex determination may be involved in mammalian glucose homeostasis. This novel pathway could be involved in the β-cell dysfunction seen in type 2 diabetes mellitus. Since Calpain-10/NIDDM1 is known to interact with a gene that is near where human FEM1B localizes, to increase susceptibility to type 2 diabetes[24], whether FEM1B could be the responsible interacting gene becomes a pertinent question Although the mechanism of this regulation by Fem1b remains to be established, this finding strengthens evidence that a genetic pathway

homologous to nematode sex determination may be involved in mammalian glucose homeostasis and promises to offer insight into novel genes and processes as potential candidates in the pathogenesis of diabetes mellitus.

Multi-modality analysis has been emerging as a major trend in the remote sensing, computer visualization, and biomedical image fusion. Fusing biomedical images is a very challenging problem because of the possible vast content change and non-uniform distributed intensities of the involved images. The new algorithm presented in this chapter, which consists of the Adaptive Exploratory Algorithm for the control point detection and heuristic optimization fusion, is reliable and time efficient. The new approach has achieved an excellent result by giving the visualization of fundus image with a complete angiogram overlay. By locking the multi-sensor images in one place, the algorithm allows ophthalmologists to match the same eye over time to get a sense of disease progress and pinpoint surgical tools to increase accuracy and speed of the surgery. The new algorithm can be easily expanded to human or animals' 3D eye, brain, or body image feature extraction, registration, and fusion. Many biomedical registration and fusion methods are still primarily used for research activity [19]. Very few of them have been developed into the integrated user-friendly computer software. The eventual aim is to developing and distributing the advanced and easy-to-use software which is suitable for various clinical environments. This plan requires intensive user interface developing work, which allows users adjusting a few threshold parameters if necessary. The user interface must be stable, simple and informative for every day clinical routine, because most of the end-users do not have much knowledge about algorithms and computer programs. Working closely with clinicians is extremely important when applying the new methods to practical clinic applications. As computation speed has been dramatically increased, real-time live ophthalmic image processing [25] will be used to handle larger and larger volumes of data in short periods. The data transmission rate, image size, higher resolution pixels, and many other issues will inevitably stress such live imaging fusion systems. The algorithms presented in this book have potential ability to handle those challenges. The presented method is a promising step towards useful clinical tools for retinopathy diagnosis, and thus forms a good foundation for further development.

6. References

[1] Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. "The Wisconsin Epidemiologic Study of Diabetic Retiopathy, II: prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years." *Arch Ophthalmol;* 102: 520-526. 1984

[2] Ryan, M; "Diabetes The Leading Cause of New Cases of Blindness"; Ohio Department of Health, Public Affairs; November 2003;

[3] Brissova, M., M. Shiota, W. E. Nicholson, M. Gannon, S. M. Knobel, D. W. Piston, C. V. Wright, and A. C. Powers. "Reduction in pancreatic transcription factor PDX-1 impairs glucose-stimulated insulin secretion"; *J Biol. Chem.* 277:P11225–11232. 2002

[4] KELLEY, D. E., and B. H. GOODPASTER. Effects of exercise on glucose homeostasis in Type 2 diabetes mellitus. *Med. Sci. Sports Exerc.*, Vol. 33, No. 6, Suppl., p S495–S501.

[5] J. MAYFIELD, "Diagnosis and Classification of Diabetes Mellitus: New Criteria"; Bowen Research Center, Indiana University, Indianapolis, Indiana

[6] Cox, N. J., M. Frigge, D. L. Nicolae, P. Concannon, C. L. Hanis, G. I. Bell, and A. Kong. "Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans"; *Nat. Genet.* 21: 213–215.1999

[7] Ventura-Holman, T., N. B. Hander, and J. F. Maher. "The human FEM1B gene maps to chromosome 15q22 and is excluded as the gene for Bardet-Biedl syndrome, type 4": *Am. J Med. Sci.* 319:268–270. 2000

[8] D. Lu, Tereza Ventura-Holman, J. Li, Robert W. McMurray, Jose S. Subauste, and Joseph F. Maher. "Abnormal Glucose Homeostasis and Pancreatic Islet Function in Mice with Inactivation of the Fem1b Gene". *Molecular and Cellular Biology.* 25(15): 6570–6577;2005

[9] Ventura-Holman, T., M. F. Seldin, W. Li, and J. F. Maher. "The murine Fem1 gene family: homologs of the Caenorhabditis elegans sex-determination protein FEM-1"; *Genomics* 54:221–230. 1998

[10] Chan, S. L., K. O. Tan, L. Zhang, K. S. Yee, F. Ronca, M. Y. Chan, and V. C. Yu. "F1Aα, a death receptor-binding protein homologous to the Caenorhabditis elegans sex-determining protein, FEM-1, is a caspase substrate that mediates apoptosis". *J Biol. Chem.* 274:32461–32468. 1999

[11] J. Claverie; C. notredame; "Bioinformatics For Dummies, 2nd Edition"; ISBN: 978-0-470-08985-9; Wiley Publishing, Inc; 2007

[12] Needleman, Saul B.; and Wunsch, Christian D. "A general method applicable to the search for similarities in the amino acid sequence of two proteins"; *Journal of Molecular Biology*; V48, I3: P443–53; 1970.

[13] Smith, Temple; and Waterman, Michael. "Identification of Common Molecular Subsequences"; *Journal of Molecular Biology;* V147: P195–197. 1981

[14] Denninghoff KR, Smith MH, Hillman L. "Retinal imaging techniques in diabetes"; *Diabetes Technol Ther.* V2:111-113. 2002

[15] Schweitzer D, Hammer M, Kraft J, Thamm E, Konigsdorffer E, Strobel J. In vivo measurement of the oxygen saturation of retinal vessels in healthy volunteers. Transactions on Biomdeical Engineering. 46: 1454-1465. 1999

[16] Yoneya A, Saito T, Nishiyama Y, et al. "Retinal oxygen saturation levels in patients with central retinal vein occlusion". *Ophthalmology*;109:1521-1526. 2002

[17] J. Beach, J. Ning, B. Khoobehi; "Oxygen Saturation in Optic Neve Head Structures by Hyperspectral Image Analysis"; Current Eye Research; V32, P161-170, 2007.

[18] B. Khoobehi, J. Beach, H. Kawano; "Hyperspectral imaging for measurement of oxygen saturation in the optic nerve head"; Invest Ophthalmol Vis Sci. V45, P1464–1472; 2004

[19] H. Cao; "Automated Fusion of Multi-Modality Biomedical Images"; ISBN: 363906622;VDM, Verlag Dr. Muller Publisher, August 2008.

[20] J.F. Canny; "A computational approach to edge detection"; IEEE Transactions on Pattern Analysis and Machine Intelligence; V8, P 679-698; 1986.

[21] H. Cao, B. Khoobehi, S. S. Iyengar; "Automated Optic Nerve Head Image Fusion of Nonhuman Primate Eyes Using Heuristic Optimization Algorithm"; 5th IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB 2008); Sun Valley, Idaho, USA; September 15-17, 2008; P 228-232

[22] Z. Millwala; "A dual-state approach to dental image registration"; Masters Thesis; *Lane Department of Computer Science and Electrical Engineering; West Virginia University*; 2004

[23] Denninghoff KR, Smith MH, Hillman L. "Retinal imaging techniques in diabetes"; *Diabetes Technol Ther.* V2:111-113. 2002

[24] Schweitzer D, Hammer M, Kraft J, Thamm E, Konigsdorffer E, Strobel J. In vivo measurement of the oxygen saturation of retinal vessels in healthy volunteers. Transactions on Biomdeical Engineering. 46: 1454-1465. 1999

[25] Yoneya A, Saito T, Nishiyama Y, et al. "Retinal oxygen saturation levels in patients with central retinal vein occlusion". *Ophthalmology*;V109: P1521-1526. 2002

Biomedical Science, Engineering and Technology

Edited by Prof. Dhanjoo N. Ghista

This innovative book integrates the disciplines of biomedical science, biomedical engineering, biotechnology, physiological engineering, and hospital management technology. Herein, Biomedical science covers topics on disease pathways, models and treatment mechanisms, and the roles of red palm oil and phytomedicinal plants in reducing HIV and diabetes complications by enhancing antioxidant activity. Biomedical engineering coves topics of biomaterials (biodegradable polymers and magnetic nanomaterials), coronary stents, contact lenses, modelling of flows through tubes of varying cross-section, heart rate variability analysis of diabetic neuropathy, and EEG analysis in brain function assessment. Biotechnology covers the topics of hydrophobic interaction chromatography, protein scaffolds engineering, liposomes for construction of vaccines, induced pluripotent stem cells to fix genetic diseases by regenerative approaches, polymeric drug conjugates for improving the efficacy of anticancer drugs, and genetic modification of animals for agricultural use. Physiological engineering deals with mathematical modelling of physiological (cardiac, lung ventilation, glucose regulation) systems and formulation of indices for medical assessment (such as cardiac contractility, lung disease status, and diabetes risk). Finally, Hospital management science and technology involves the application of both biomedical engineering and industrial engineering for cost-effective operation of a hospital.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH

open science | open minds

www.intechopen.com