Facultade de Informática

Departamento de Computación

PhD Thesis

Advancing the diagnosis of Dry Eye Syndrome: development of dynamic, automated tear film Break-Up assessment

Lucía Ramos García

November 2014

PhD Advisors:

Dra. Noelia Barreira Rodríguez

Dr. Antonio Mosquera González

Dra. Noelia Barreira Rodríguez y Dr. Antonio Mosquera González

DECLARAN:

Que la memoria titulada "Advancing the diagnosis of Dry Eye Syndrome: development of dynamic, automated tear film Break-Up assessment", ha sido realizada por Dª. Lucía Ramos García bajo nuestra dirección en el Departamento de Computación de la Universidade da Coruña y concluye la Tesis que presenta para optar al grado de Doctor con la mención de Doctor Internacional.

A Coruña, 25 de noviembre de 2014

Fdo: Noelia Barreira Rodríguez
Directora de la Tesis doctoral
Departamento de Computación
Facultad de Informática
Universidade da Coruña

Fdo: Antonio Mosquera González
Director de la Tesis doctoral
Departamento de Electrónica y Computación
Escola Técnica Superior de Enxeñaría
Universidade de Santiago de Compostela

To my family

Agradecimientos

En primer lugar me gustaría expresar mi agradecimiento a mis directores de tesis por toda la ayuda que me han prestado durante esta etapa, por la paciencia que han tenido conmigo y por todo lo que he aprendido junto a ellos. Noelia, moitas gracias por tódalas boas propostas, pola dedicación coa que revisas todas as veces que faga falta, por estar sempre disposta a pesar de andar a tantas cousas, por ser tan positiva e facer que saia motivada cada vez que paso polo teu despacho e por coidar tanto do grupo. Antonio, muchas gracias por todas tus indicaciones y buenas ideas, por transmitirme calma y confianza, por tu optimismo y sentido del humor.

Tambien me gustaría dar las gracias a Penedo por darme la oportunidad de entrar en el VARPA, por abrirme una ventana al mundo de la investigación, por pensar en mi para ofrecerme nuevas posibilidades y por todo lo que me ha aportado formar parte del grupo.

Gracias a mis compañeros del VARPA, porque aunque siempre bromee con que estoy exiliada en el CITIC, el buen ambiente del grupo llega hasta aquí arriba. Gracias por todas las risas y buenos momentos compartidos en cenas, comidas, congresos.. y también porque sé que puedo contar con vuestra colaboración cada vez que lo necesito.

Gracias también a los compañeros de tupper del CITIC, que hacen que la hora de la comida sean un respiro en el que darnos a la creatividad, porque solo con vosotros pueden surgir todos esos inventos descabellados que nos harán ricos en el futuro, o que al menos nos permiten desconectar un rato en esos días que se nos pasan encerrados en un edificio.

Este trabajo no habría sido posible sin la colaboración del Servizo de Optometría de la Universidade de Santiago de Compostela. Gracias por propor-

cionar los conjuntos de datos con su laborioso etiquetado manual, por ofrecer vuestro conocimiento desde el punto de vista clínico y por valorar tan positivamente los resultados obtenidos.

I would like to thank to Dr Katherine Oliver for allowing me to enjoy a stay at Glasgow Caledonian University (Scotland) and work in a renowned group in dry eye research. Thanks to Eilidh Martin and the whole group for their warm welcome and their help in my research. Definitely, my stay in Scotland was an unforgettable experience where I had the opportunity to meet great people and visit amazing places.

No podría estar más agradecida por todos los amigos y amigas que tengo. Por los de siempre y por muchas personas nuevas que han ido apareciendo a lo largo de esta etapa. Gracias por alegrarme la vida, por animarme en todo, por estar siempre ahí, por entenderme cada vez que rechazo vuestras propuestas diciendo que estoy muy liada y por seguir proponiendo grandes planes. Por todos los buenos momentos compartidos y por los que seguro que vendrán.

Gracias a mi hermano, mis abuelos, mis tíos y mis primos. Gracias a mis padres por quererme y apoyarme siempre.

"And once the storm is over,
you won't remember how you made it through, how you managed to survive.
You won't even be sure whether the storm is really over.
But one thing is certain. When you come out of the storm,
you won't be the same person who walked in.
That's what this storm's all about."
Haruki Murakami

Resumen

El síndrome de ojo seco es un trastorno común de la película lagrimal que afecta a un sector significativo de la población, impactando en la calidad de vida. El diagnóstico de esta enfermedad es difícil debido a su etiología multifactorial, por lo que hay varias pruebas clínicas para evaluar diferentes aspectos de la película lagrimal. Una de las pruebas empleadas habitualmente es el test de BUT, que consiste en medir el tiempo transcurrido desde el último parpadeo hasta la ruptura de la película lagrimal, representada por la aparición de áreas oscuras que corresponden al adelgazamiento de la película lagrimal en la superficie ocular. Además del tiempo, hay varias características de la ruptura como la zona, la forma, el tamaño y la evolución, que podrían afectar a la severidad del síndrome de ojo seco. Sin embargo, el test de BUT presenta una baja repetibilidad debido principalmente a la apreciación subjetiva de los puntos oscuros, las diferencias entre expertos y la variabilidad de la película lagrimal. Además, la caracterización a mano de las zonas de ruptura es una tarea tediosa que consume mucho tiempo. La automatización del análisis de la ruptura reduciría su carácter subjetivo, permitiendo una evaluación más precisa de la película lagrimal.

Este trabajo presenta una metodología novel para una evaluación de la ruptura de la película lagrimal totalmente automática. Este estudio permite un análisis cuantitativo y cualitativo de la inestabilidad de la película lagrimal como una extensión de la medida de BUT, que se centra solo en el tiempo. Esta metodología proporciona resultados de precisión en los mismos rangos que entre los propios expertos. Así, la evaluación automática de la ruptura ahorra tiempo a los expertos proporcionando resultados imparciales que no están afectados por factores subjetivos.

Resumo

O síndrome de ollo seco é un trastorno común da película lacrimal que afecta a un sector significativo da poboación, impactando na calidade de vida. A diagnose desta enfermidade é difícil debido a súa etioloxía multifactorial, polo que hai varias probas clínicas para avaliar diferentes aspectos da película lacrimal. Unha das probas empregadas habitualmente é o test de BUT (Break-Up Time), que consiste en medir o tempo transcorrido dende o último pestanexo ata a ruptura da película lacrimal, representada pola aparición de áreas escuras que corresponden ó adelgazamento da película lacrimal na superficie ocular. Ademais do tempo, hai varias características da ruptura como a zona, forma, tamaño e evolución, que poderían afectar á severidade do síndrome de ollo seco. Sen embargo, o test de BUT presenta unha baixa repetibilidade debido principalmente á apreciación subxectiva dos puntos escuros, ás diferencias entre expertos e á variabilidade da película lacrimal. Ademais, a caracterización á man das zonas de ruptura é unha tarefa tediosa que consume moito tempo. A automatización da análise da ruptura reduciría o seu carácter subxectivo, permitindo unha avaliación máis precisa da película lacrimal.

Este traballo presenta unha metodoloxía novel para unha avaliación da ruptura da película lacrimal totalmente automática. Este estudo permite unha análise cualitativa e cuantitativa da inestabilidade da película lacrimal como unha extensión da medida de BUT, que se centra só no tempo. Esta metodoloxía proporciona resultados de precisión nos mesmos rangos que entre os propios expertos. Deste xeito, a avaliación automática da ruptura aforra tempo ós expertos proporcionando resultados imparciais que non están afectados por factores subxectivos.

Summary

Dry Eye Syndrome (DES) is a common disorder of the tear film which affects a significant sector of the population, impacting on quality of life. The diagnosis of this condition is difficult due to its multifactorial etiology, so there are a wide number of tests to evaluate different aspects of the tear film. Among the different tests available, the Break-up Time test (BUT) is a widely used measure to evaluate the quality and stability of the tear film on the ocular surface. It consists in measuring the time elapsed since the last blink until the tear film break-up, that is, the appearance of dark areas related to a thinning of the tear film on the surface of the eye. Besides the time, there are several break-up features such as the location, shape, size and dynamics of the break-up areas, which could affect to dry eye severity. However, the BUT test is affected by low repeatability mainly due to a subjective appreciation of the dark spots, the differences among the experts, and the variability of the tear film. Furthermore, the characterization by hand of break-up areas is a tedious and time consuming task. The automation of the break-up assessment would reduce its subjective character, allowing a more accurate evaluation of tear film stability.

This work presents a novel methodology for a fully automatic assessment of the tear film break-up. The proposed methodology allows a quantitative, qualitative analysis of tear film instability, as an extension of BUT measurement, which is focused only on time. It provides accuracy results that are in the same ranges as the experts themselves. Therefore, the automated break-up assessment saves time for experts providing unbiased results which are not affected by subjective factors.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The eye is the most valuable, sensitive sense organ and is the basis of the sight. The eye's function consists in transforming the light energy into electrical signals that are sent to the brain via the optic nerve. It represents a window to the world, being responsible for four fifths of all the information our brain receives [Acharya and Suri, 2008]. The eyeball is a slightly asymmetrical globe composed of three tunics or layers and a refracting media called the humours. The first and most inner tunic is the *retina*; the second one is composed of the *choroid*, the *ciliary body* and the *iris*; and, finally, the third and external tunic is formed by the *sclera* and the *cornea* (see Figure 1.1).

The cornea is a transparent structure found in the very front of the eye that helps to focus incoming light. There are more nerve endings in the cornea than anywhere else in the body, so it is extremely sensitive. To remain healthy, the cells of the corneal epithelium must be kept moist. This need is met by the elaboration of a thin layer of lubricating substances known as the precorneal tear film. The lacrimal glands, located in the upper, outer portion of each orbit, secrete the tears, which flow through the main excretory ducts into the space between the eyeball and lids. When a blink occurs, the lacrimal fluid is spread, cleaning and lubricating the surface of the eyes [Hughes, 1991].

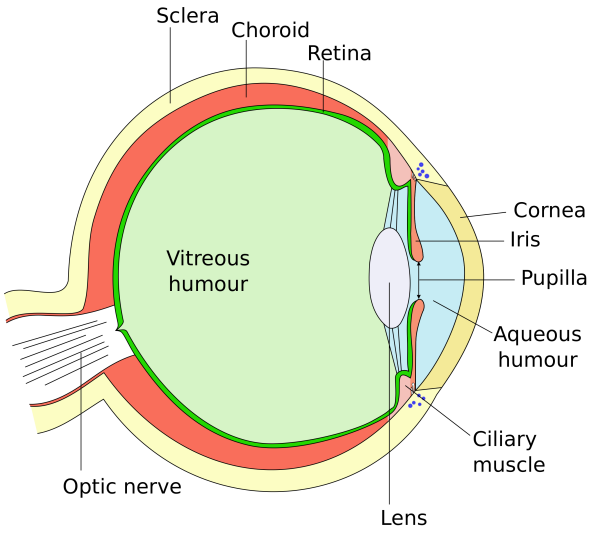


Figure 1.1: Schematic diagram of the human eye.

1.1 The Tear Film

The tear film is a complex and dynamic structure of lipids, proteins and mucins covering the anterior surface of the cornea. It is an essential component of the eye that plays important functions to guarantee a high optical quality vision [Acharya and Suri, 2008]:

- Lubrication function. Tear film minimizes the friction between eyelid margins and palpebral conjunctiva during the blinking.
- Visual function. The tear film fills the small irregularities in the corneal epithelium providing a smooth, regular optical surface to guaranties a high optical quality vision.
- Cleaning function. Tear film accepts the desquamated epithelial cells, debris, etc. from the epithelium, and, together with blinking, remove them from the surface of the eye.
- Nutritive function. Corneal surface is avascular to guaranty its transparency, so the nutrition is driven by the tear film. Oxygen from the

ambient air dissolves in the tear fluid and is transferred to the corneal epithelium.

- Antimicrobial function. The tear film is the first line of defense against ocular surface infection. It contains proteins, such as lysozyme or lactoferrin, that inhibit microbiological contamination.

Classically, the tear film is described as a trilaminar structure comprising a thin anterior lipid layer, an intermediate aqueous layer, and an innermost mucous layer [Larke, 1997], as depicted in figure 1.2.

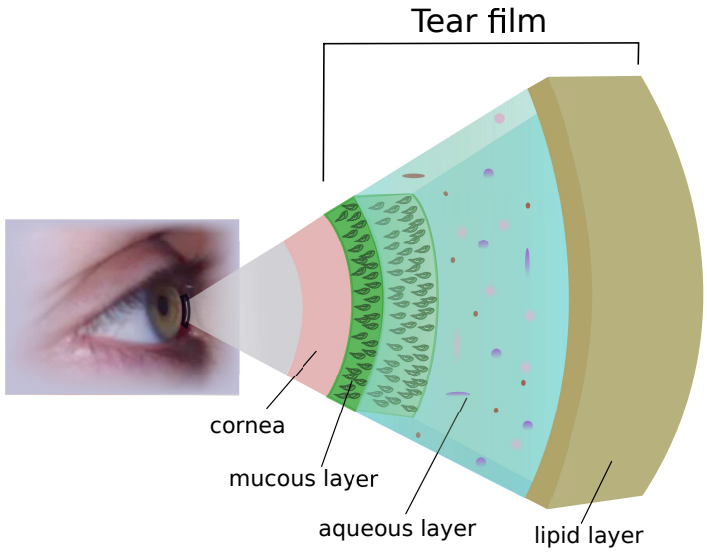


Figure 1.2: Trilaminar structure of the tear film composed by an outer lipid layer, an intermediate aqueous layer, and an inner mucous layer.

Lipid layer $(0.05 - 0.1\mu l)$ is the outermost and thinnest layer of the tear film and it is mainly secreted by the meibomian glands, embedded in the upper and lower tarsal plates [Nichols et al., 2004]. Normal lipid layer is divided into two layers: anterior lipid layer formed by non-polar lipids, and posterior lipid layer formed by high-polar lipids [Korb, 2002]. The main function of this layer is the reduction of evaporation from the

aqueous phase. Furthermore, the non-polar structure is important for preventing the ocular surface contamination, which could disrupt the tear film.

Aqueous layer $(6.5 - 7.5\mu l)$ comprises around 98% of the total thickness in the tear film [Larke, 1997]. It is a complex dilute solution of both inorganic electrolytes and low and high molecular weight organic substances. This phase is mainly secreted by lacrimal glands located in the superior temporal angle of the orbit. This phase contains many ions and molecules such as electrolytes, hydrogen ions, proteins, enzymes, and metabolites that provides the proper functions of the tear film.

Mucous layer $(0.02 - 0.04\mu l)$ is the innermost layer, representing the 0.2% of the whole tear film. It is secreted by globet cells sited in the conjunctiva. The main functions are the maintenance of the surface tension and the lubrication of the cornea, allowing the lids to slide with minimal friction during the blinks and wetting the ocular surface to maintain an stable tear film [Larke, 1997].

Each of these layers has a specific role in the formation and stability of the tear film. The quality and thickness of each layer, as well as their adequate interaction, are important in order to have a stable tear film. Abnormalities in any of the layers can cause tear dysfunction problems.

1.2 Dry Eye Syndrome

The lipid layer plays a major role in retarding the evaporation of the tear film during the inter-blink period, and consequently, a deficit of this layer can cause complications, such as the Dry Eye Syndrome (DES) [Graig and Tomlinson, 2007]. The DES is a multifactorial disease of the ocular surface,

which affects a significant percentage of the population, and worsens with age [Javadi and Feizi, 2011; Lemp, 2008; Lowther, 1977]. The prevalence of this syndrome has been increasing in recent years, affecting up to 10-15% of normal population, and 18-30% of contact lenses users. Several factors, such as adverse environmental conditions, use of certain medications, or visual tasks that reduce blink rate, have contributed to that increment [Fenga et al., 2008; García-Resúa et al., 2005]. Blink rate is an important factor in tear film stability, since if it is reduced, the ocular surface is exposed to water loss, increasing the tear film evaporation [Abelson et al., 2002]. Contact lenses wear also may cause dry eye, since it disrupts the tear film [Kastelan et al., 2013].

According to the main etiological causes of the disease, DES can be classified in two main categories, Aqueous Tear-Deficient Dry Eye and Evaporative Dry Eye [Lemp et al., 2007; Lemp, 1995], which are not mutually exclusive.

Aqueous Tear-Deficient Dry Eye (ADDE) is due to a failure of lacrimal tear secretion. This causes tear hyperosmolarity and stimulates a cascade of inflammatory events. ADDE has two major subclasses, Sjogren Syndrome Dry Eye and Non-Sjogren Syndrome Dry Eye. Sjogren Syndrome is an exocrinopathy in which the exocrine glands (such as lacrimal and salivary glands) are targeted by an autoimmune process. On the other hand, Non-Sjogren Syndrome Dry Eye does not present those systemic autoimmune features. The most common form of Non-Sjogren Syndrome Dry Eye is age-related dry eye, although other factors can contribute to ADDE, such as lacrimal gland infiltration, sarcoidosis, lymphoma, obstruction of the lacrimal gland ducts, or reflex hyposecretion [Lemp et al., 2007; Lemp, 1995].

Evaporative Dry Eye (EDE) is caused by excessive water loss from the exposed ocular surface in the presence of normal lacrimal secretory function. The volume and composition of the lacrimal fluid are adequate, whereas

tear abnormality is created by other periocular disease, usually leading to increase tear evaporation. This is the type of Dry Eye most commonly found in young to middle-aged people, and related to ambient conditions and/or contact lenses wear [Lemp et al., 2007; Lemp, 1995]. The current working conditions, such as computer use, have increased the proportion of people with EDE [Lemp et al., 2007]. The main cause of EDE is the Meibomian Gland Dysfunction [Bron et al., 2004]. Meibomian Glands, embedded in the upper and lower tarsal plates are responsible of lipid secretion, essential to retard tear film evaporation [Bron et al., 2004].

DES affects the quality of life in different ways. Pain and irritative symptoms as blurry vision, grittiness, redness, soreness, foggy sensation, ocular fatigue, and lost of contrast sensitivity disrupt negatively with the welfare of the patient [Tutt et al., 2000]. Thus, DES limits and degrades visual performance impacting to workplace productivity and daily activities, such as reading, using a computer, watching television, or driving [Nichols et al., 1999]. Consequently, patients with DES are about three times more likely to report problems with common activities [Miljanovic et al., 2007]. DES is also associated with contact lenses intolerance and discontinuation of contact lenses wear. It can affect to refractive surgery outcomes and may be associated with increased risk of infection and complications with ocular surgery. Severe dry eye may lead to ocular damage [Stevenson et al., 2012].

The high prevalence among older people, along with the aging of the population make DES a public health problem with the potential for a high economic burden. DES supposes an economic impact derived from costs due to health care system utilization, including health care professional visits, non pharmacological therapies, pharmacological treatments, surgical procedures, and other therapeutics, such as humidifiers [Reddy et al., 2004]. Additionally, costs derived from daily capacity can be affected, such as lost working time and

productivity, or alteration in work type or environment. Intangible costs include decreased leisure time, impaired physical functioning and quality of life, as well as impact on social interactions, mental, and general health [Miljanovic et al., 2007].

The diagnosis of DES cannot be based only on symptoms, since most of them can be also caused by other conditions. The composition and behavior of the tear film provides crucial indicators, so the tear film assessment is essential for DES characterization.

1.3 Tear Film Assessment

The diagnosis of DES is very difficult to achieve due to its multifactorial etiology [Khanal, 2008]. There are a wide number of tests to evaluate different aspects of the tear film, but most of them have high variability and low repeatability [Lemp et al., 2007]. This determines that clinicians recommend a large number of tests to obtain a clear diagnosis.

The different dry eye diagnostic tests can be grouped in several categories, depending on which tear film parameters are evaluated. Figure 1.3 illustrates a taxonomy for these tests. Thus, there are dry eye questionnaires, represented in red, which are used for a preliminary evaluation; the examination by the biomicroscope, in purple, which is as an essential part within the protocol; clinical tests which are divided into two groups, quantitative tests, in dark green, for assessing the tear secretion, and qualitative tests, in light green, for assessing the tear film stability; the analysis of the osmolarity, in orange, also provides a powerful indicator for DES diagnosis; and, finally, the laboratory tests, in blue, which are focused on the tear film composition.

The clinical tools most commonly used to grade DES severity are the Schirmer test, the tear film Break-Up Time (BUT) test, the meibomian se-

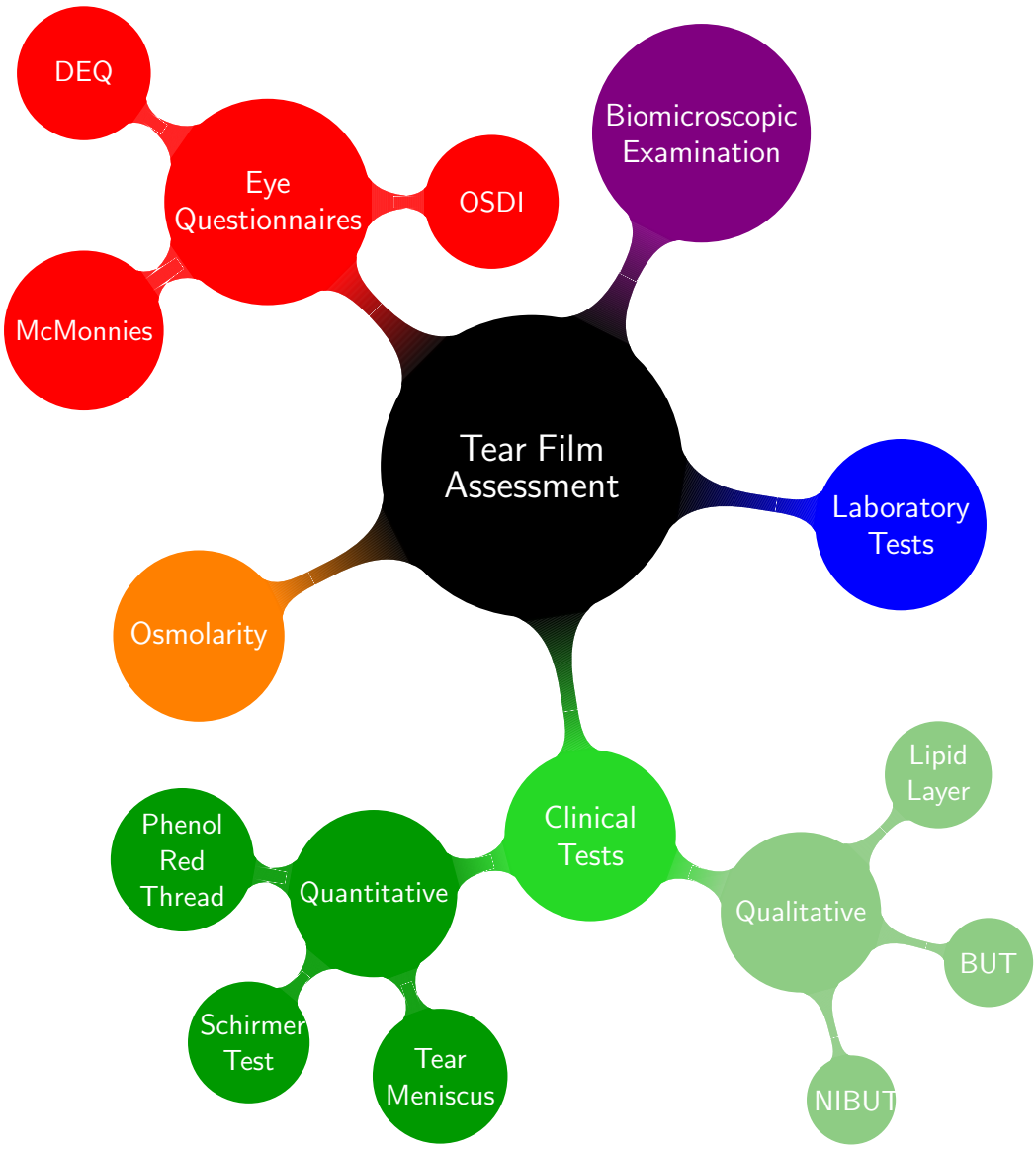


Figure 1.3: Taxonomy of tear film assessment tests

cretion scoring, the fluorescein green staining of the cornea and conjunctiva, or symptom questionnaires, such as McMonnies and Ocular Surface Disease Index (OSDI). Recently, measurement of tear film osmolarity has been suggested as a gold standard for the diagnosis of DES since high tear film osmolarity is thought to be a key mechanism in symptoms and ocular surface damage in dry eye. The clinicians recommend to perform a battery of tests and combine the obtained results for getting a diagnosis based on different factors.

In daily clinical practice is often used a short protocol with some common tests for the study of the anterior segment of the eye that includes the analysis of the conjunctiva, cornea, and tear film. More specific tests are used when there are evidences of severe DES or when the former tests do not achieve definitive results.

1.3.1 Dry Eye Questionnaires

There are several questionnaires for a primary evaluation of ocular comfort. Dry eye is a symptomatic disease, so the presence and status of symptoms needs to be ascertained. Symptoms vary in severity according to the state of instability of the tear film and damage of the ocular surface [Korb, 2002]. One problem that can arise is the subjective interpretation of symptoms by the patient. The standardized questionnaires provide an objective symptom evaluation and guarantee comparisons [Simpson et al., 2008]. There are available several questionnaires, but the most used are McMonnies Questionnaire [Gothwal et al., 2010; McMonnies and Ho, 1987a,b], Ocular Surface Disease Index (OSDI) questionnaire [Dougherty et al., 2011], and Dry Eye Questionnaire (DEQ) [Chalmers et al., 2010]. Although these questionnaires alone may not be a reliable measure of severity, they are useful indicators. Thus, symptoms such as pain, blurred vision, foreign body sensation, ocular discomfort, itching, tearing, and photophobia are examinated in several situations of ev-

eryday life and with different conditions of humidity and lighting. Patients are given four options to answer, i.e. never, rarely, sometimes, often/all the times, and are asked to report the frequency of each symptom. Patients reporting one or more symptoms as often/all the times are considered as symptomatic patients [Vashisht and Singh, 2011]. Figure 1.4[1] shows the questions of OSDI questionnaire while Fig. 1.5[1] shows how the experts evaluate the OSDI score and get a value for dry eye severity .

1.3.2 Biomicroscopic Examination

Biomicroscopic examination is an essential part within the protocol of eye examination. It consists of a biomicroscope designed to view anterior eye with high magnification, and a slit lamp to change the illumination technique, as shown in Fig. 1.6 (left)[1]. The biomicroscope is essential for grading ocular surface staining, that is an indicative sign of epithelial damage. There are various stains available, being sodium fluorescein and lissamine green the most used. Sodium fluorescein is a dye indicated to observe corneal epithelium. Under biomicroscopic observation with a cobalt blue and yellow filters, potential corneal epithelium damage can be seen as bright dots, as shown in Fig. 1.6 (right)[1].

1.3.3 Clinical Tests

Clinical tests are those that can be performed in clinical settings, as part of a routine eye care examination. Classically, these tests have been divided into two groups, quantitative and qualitative tests. Quantitative tests assess quantity of tear film and are related to tear secretion, whereas qualitative tests assess the stability of tear film [Lemp, 1995].

[1]Image courtesy of the Servizo de Optometría of the Universidade de Santiago de Compostela.

Ocular Surface Disease Index© (OSDI©)[2]

Ask your patients the following 12 questions, and circle the number in the box that best represents each answer. Then, fill in boxes A, B, C, D, and E according to the instructions beside each.

Have you experienced any of the following *during the last week*?	All of the time	Most of the time	Half of the time	Some of the time	None of the time
1. Eyes that are sensitive to light? . .	4	3	2	1	0
2. Eyes that feel gritty?	4	3	2	1	0
3. Painful or sore eyes?	4	3	2	1	0
4. Blurred vision?	4	3	2	1	0
5. Poor vision?	4	3	2	1	0

Subtotal score for answers 1 to 5 (A)

Have problems with your eyes limited you in performing any of the following *during the last week*?	All of the time	Most of the time	Half of the time	Some of the time	None of the time	N/A
6. Reading?	4	3	2	1	0	N/A
7. Driving at night?	4	3	2	1	0	N/A
8. Working with a computer or bank machine (ATM)?	4	3	2	1	0	N/A
9. Watching TV?	4	3	2	1	0	N/A

Subtotal score for answers 6 to 9 (B)

Have your eyes felt uncomfortable in any of the following situations *during the last week*?	All of the time	Most of the time	Half of the time	Some of the time	None of the time	N/A
10. Windy conditions?	4	3	2	1	0	N/A
11. Places or areas with low humidity (very dry)?	4	3	2	1	0	N/A
12. Areas that are air conditioned? . . .	4	3	2	1	0	N/A

Subtotal score for answers 10 to 12 (C)

Add subtotals A, B, and C to obtain D
(D = sum of scores for all questions answered) (D)

Total number of questions answered
(do not include questions answered N/A) (E)

Please turn over the questionnaire to calculate the patient's final OSDI© score.

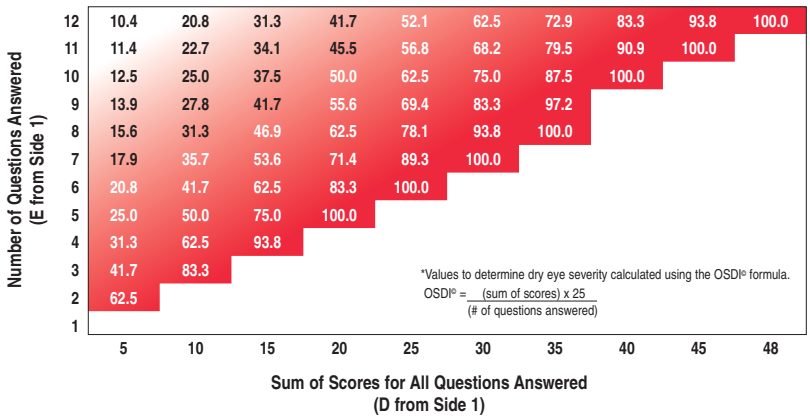
Figure 1.4: Ocular Surface Disease Index (OSDI) Questionnaire.

Evaluating the OSDI© Score[1]

The OSDI© is assessed on a scale of 0 to 100, with higher scores representing greater disability. The index demonstrates sensitivity and specificity in distinguishing between normal subjects and patients with dry eye disease. The OSDI© is a valid and reliable instrument for measuring dry eye disease (normal, mild to moderate, and severe) and effect on vision-related function.

Assessing Your Patient's Dry Eye Disease[1, 2]

Use your answers D and E from side 1 to compare the sum of scores for all questions answered (D) and the number of questions answered (E) with the chart below.* Find where your patient's score would fall. Match the corresponding shade of red to the key below to determine whether your patient's score indicates normal, mild, moderate, or severe dry eye disease.



Patient's Name: _____ Date: _____

How long has the patient experienced dry eye disease? _____

Eye Care Professional's Comments: _____

1. Data on file, Allergan, Inc.
2. Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL. Reliability and validity of the Ocular Surface Disease Index. *Arch Ophthalmol.* 2000;118:615-621

Figure 1.5: Evaluation of Ocular Surface Dry Eye Illness (OSDI) Questionnaire. The OSDI is assessed on a scale of 0 to 100, getting an index of dry eye severity.

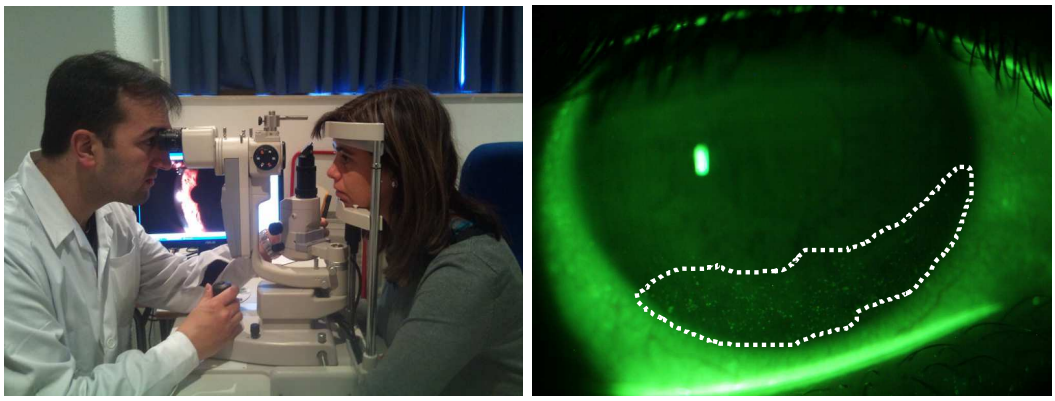


Figure 1.6: Biomicroscopic examination. Left: eye examination under biomicroscope with a slit lamp. Right: Appearance of anterior eye dyed with fluorescein under biomicroscopic observation with cobalt blue and yellow filters. Light spots within the dotted line indicate damage of corneal epithelium surface.

Quantitative Tear Film Tests are related to the lacrimal gland secretion. Defective lacrimal function is usually demonstrated by showing reduced aqueous tear volume and tear flow. Some of the more used quantitative tests are:

Tear Meniscus Assessment. It measures the tear reservoir along the low lid. The tear meniscus supposes the $75 - 90\%$ of tear volume, so it is an useful indicator [García-Resúa et al., 2009]. Thus, the Tear Meniscus Height (TMH) is the most used parameter for tear volume examination. It can be assessed in clinical settings but it is necessary a biomicroscope with high magnification to obtain enough resolution. With the aid of a graticule eyepiece it is possible to measure THM, from the lid to the top of the meniscus (see Fig. 1.7[1]). The cut-off point between normal and dry eye is less than 0.1 *mm*. Tearscope Plus, a device designed to evaluate the anterior lipid layer also can be used to evaluate TMH (see Fig. 1.7, right[1]). The main advantage of the Tear Meniscus assessment is its non invasive nature, but the reliability of this evaluation is affected by subjective criteria of the observer.

Figure 1.7: Appearance of tear meniscus height by slit lamp biomicroscope.

The Schirmer Test. It measures the tear flow reflexively stimulated by insertion of a filter paper into the conjunctival sac (see Fig. 1.8[1]). The strip is placed at the junction of the middle and lateral on-third of the lower eye lid and the patient is told to keep the eyes closed. After 5 minutes, the strip is replaced and the length of wet strip is measured. The cut-off value to distinguish normal and dry eye subjects is established in 5 mm. Although this is a test widely used, it is affected by a wide variability [Korb, 2002]. Furthermore, this test is very uncomfortable because the patient has to bear the strip in his/her eye for several minutes.

Figure 1.8: The Schirmer Test. Left: Strips for measuring the tear production on a period of time. Right: Schirmer test in use. The strip is inserted into the lower lid and the patient keeps the eye closed for the test duration.

Phenol Red Thread Test. It consists of a thread impregnated with phenol red, which is pH sensitive and changes from yellow to red over the pH range of normal tears (see Fig. 1.9[1]). This test only needs to be hooked over the lower lid for 15 seconds and it is more comfortable than Schirmer, since it is barely noticeable for the patient. The cut-off value between controls and dry eye is 10 mm [Lemp et al., 2007].

Figure 1.9: Phenol red thread test in use.

Qualitative Tear Film Tests are related to the ability of the tear film to remain stable, which is essential to cover the anterior eye and perform its functions (optical, nutritional, antimicrobial, and cleaning). Secretion of the lacrimal gland may be normal and the volume and composition of the lacrimal fluid adequate, but there could be any tear abnormality driven by other factors that can lead to increased tear evaporation.

Lipid Layer Pattern Assessment. It consists in analyzing the lipid layer structure through the observation of the interference phenomena, since the color and shape of the observed patterns reflect the layer thickness

[Guillon, 1998; Remeseiro et al., 2013]. Thicker lipid layers ($\geq$ 90 nm) show color and wave patterns, while thinner lipid layers ($\leq$ 60nm) are more homogeneous. The Tearscope Plus, designed by Guillon [Guillon, 1998], is the instrument of choice for rapid assessment of lipid layer thickness in clinical settings. Guillon also proposed five main categories of lipid interference patterns in increasing thickness: marmoreal (open and closed meshwork), wave, amorphous, and color fringes (see Fig. 1.10[1]).

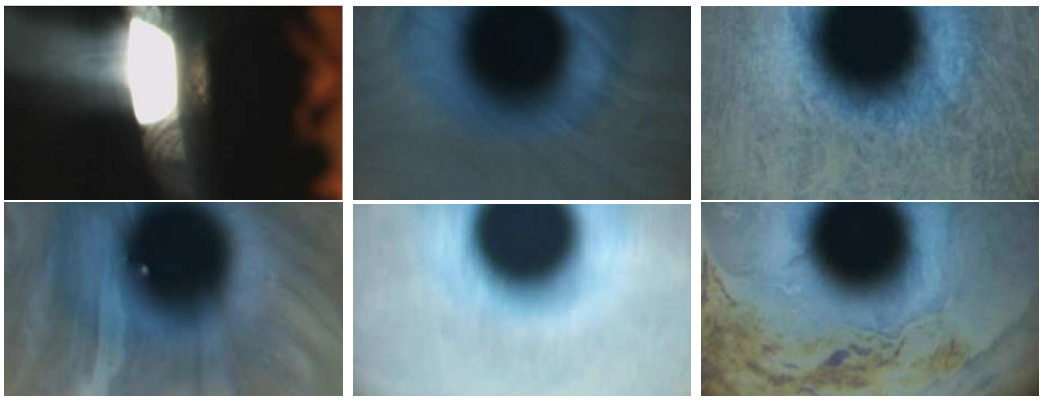


Figure 1.10: From top to bottom, left to right: Appearance of lipid layer by slit-lamp examination, open meshwork, closed meshwork, wave, amorphous, and color fringe lipid layer thickness categories.

However, the classification of the lipid layer thickness is a difficult clinical technique, especially with thinner lipid layers that lack distinct features, and is affected by the subjective interpretation of the observer.

Break-Up Time Test (BUT). It consists in measuring the time that the tear film remains stable without blinking. After fluorescein instillation, the patient is asked to keep the eye open until a sign of tear film rupture (dark spot) appears (see Fig. 1.11[1]). This is the most commonly used test of tear film stability and the most used cut-off for dry eye diagnosis is of less than 5 seconds [Abelson et al., 2002].

Not Invasive Tear Break-Up Time Test (NIBUT). This test uses a pattern directed onto the precorneal tear film for the observation of distor-

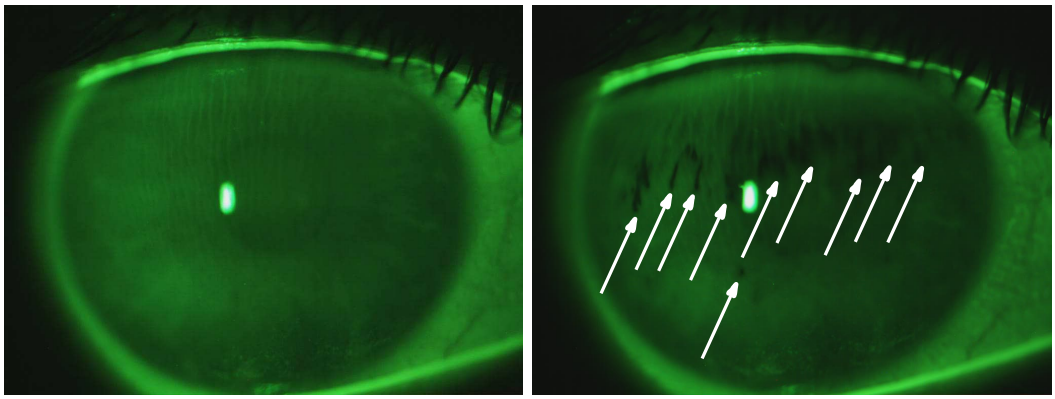


Figure 1.11: Appearance of tear film stained by fluorescein under biomicroscopic observation with cobalt blue and yellow filters. Left: Stable tear film. Right: formation of dark areas related to the tear film break-up.

tions and/or abnormalities in the image (see Fig. 1.12[1]). The appearance of variations on the grid denotes the break-up. The time in seconds from a blink to the first change of the grid is defined as NIBUT. This method avoids the need of fluorescein instillation and eliminates physical disturbance to the tear film, but requires more sophisticated methods to project a grid in the anterior eye [Cho, 1993]. The Tearscope Plus is commonly used for NIBUT test by utilizing accessory removable grids [Lemp et al., 2007].

1.3.4 Osmolarity of Tear Film

The osmolarity is a powerful indicator for DES diagnosis [Lemp et al., 2007]. Hyperosmolarity of tear fluid has been recognized as a common feature of all types of DES. It is considered that the Osmolarity of the normal tear film is of 302 to 6.3 $mOsm/l$, while in DES it reaches values of 325 to 340 $mOsm/l$ [Tomlinson et al., 2006]. However, the measurement of tear film osmolarity requires very expensive instruments or laboratory equipment hard to use (see Fig. 1.13[1]). These difficulties have hindered the application of this procedure to daily practice.

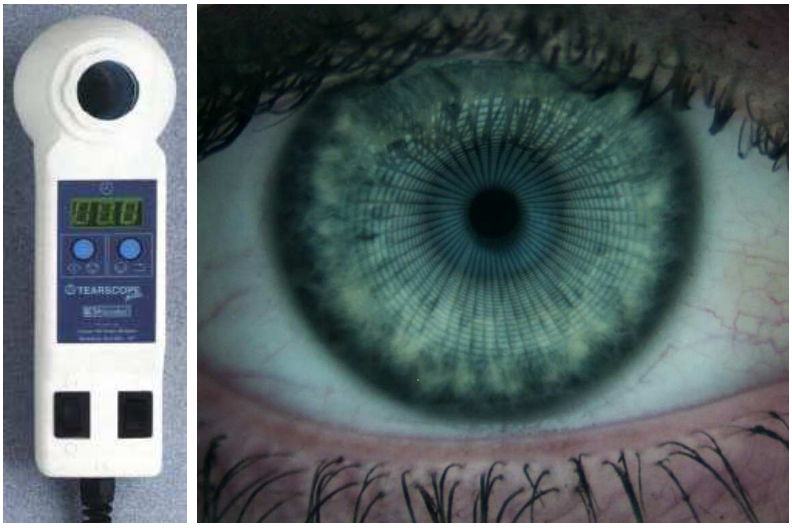


Figure 1.12: Left: Tearscope-Plus. This device projects a cylindrical source of cool white fluorescent light onto the lipid layer. Right: NIBUT test with Tearscope Plus. A grid pattern is projected onto the precorneal tear film for the observation of distortions and/or abnormalities in the image.

Figure 1.13: Left: TearLab osmometer, based on electrical conductivity. Right: TearLab in use

1.3.5 Laboratory Tests

The composition of proteins in tears plays an important role in ocular surface diseases. Several studies have demonstrated that changes in the tear protein patterns of dry eye patients compared to control patients, so proteins are useful dry eye biomarkers. The biomarkers more used are lysozyme and lactofferrin [Lemp et al., 2007]. Around of 20 to 30% of the total tear protein is made up of lysozyme, the most alkaline protein in tears. Lysozyme levels decrease with dry eye, so its estimation could be useful for the DES diagnosis. Although this test is reliable, it is an expensive and cumbersome test [Lemp, 1995]. Besides, lactoplate test determines the concentration of lactofferrin. The lactoplate uses circular discs of filter paper that are placed in the inferior conjunctival. They are placed on the agar and incubated for three days. The size of the ring is proportional to the lactoferrin concentration of the collected sample. This method, although accurate, is also too expensive to be recommended for use in clinical trials [Lemp, 1995].

1.4 Break-Up Assessment

Some of the previous tests are discarded for the usual evaluations because the equipment necessary is too specific and expensive [McGinnigle et al., 2012]. Other tests are lengthy and uncomfortable to the patient. Furthermore, some of them present excessive variability and provide results difficult to interpret. This way, the simplest and most effective tests seem to be based on analyzing the tear film stability by direct observation.

The preocular tear film does not remain stable for long periods because of the evaporation from the exposed ocular surface. Normal blinking guarantees

that a stable tear film is covering the anterior eye to maintain its adequate physiology. When blinking is prevented, the tear film breaks and dry spots appear over the cornea. A lipid film is seldom stable for long periods of time; some of the superficial lipids will migrate to the epithelium interface, contaminating the absorbed mucin layer and converting it into a hydrophobic surface.

BUT test is one of the most common tests for analyzing tear film stability. It consists of measuring the time that the tear film remains stable without blinking [Lee and Kee, 1988; Cho et al., 1992; Cho and Brown, 1993]. To perform this test, sodium fluorescein is instilled into the eye using a micropipette, and the tear film is observed with the help of cobalt-blue filter attached to a slit lamp biomicroscope, and a yellow filter to improve the visibility of the fluorescein emission [Johnson and Murphy, 2005; Elliott et al., 1998], as shown in Fig. 1.11 (left). The patient is instructed to blink three times naturally, without squeezing, in order to distribute the fluorescein over the cornea, and then, he/she maintains the eye open as long as possible [Begley et al., 2006]. The BUT is measured as the time elapsed between the last blink and the first appearance of a dark spot on the surface of the cornea, as shown in Fig. 1.11 (right), which represents the evaporation of water and the break-up of the tear film.

A low BUT measurement corresponds to a limited ocular surface wetting, and it is one of the main signs of an abnormal tear film. Classically, the cut-off value between normal and dry eye is being considered as less than 10 seconds but this cut-off value has been questioned by many authors. In this sense, some studies have found that the mean BUT for normal subjects is 7.1 seconds (range of $4.7 - 11.4$ seconds) and, for dry eye patients, 2.2 seconds (range of $0.9 - 5.2$ seconds) [Abelson et al., 2002]. On the basis of this fact, a cut-off for dry eye diagnosis of less than 5 seconds is recommended.

Although BUT reflects an invasive character, it remains the most frequently

used diagnostic test to determine tear film instability [Sweeney et al., 2008]. It is known that fluorescein can affect the measurements, but it was found that controlled pipetting of 2% fluorescein solution into the tear film instead of classic fluorescein strips, improves the repeatability of BUT measurements [Johnson and Murphy, 2005; Elliott et al., 1998]. In this sense, specific BUT strips that delivered 5 times less fluorescein than normal strips have been designed, obtaining improved reproducibility [Korb et al., 2001]. Thus, these strips can avoid the inconveniences derived by fluorescein instillation. Furthermore, although the NIBUT test exists, this is, a non-invasive version of the BUT test based on the projection of a grid on the corneal surface and the analysis of its deformations, the detection the break-up on the green dye is easier. Moreover, some grid distortions in the NIBUT test denote tear thinning instead of tear break-up so their characterization is not straightforward [Patel et al., 1985].

The BUT can be manually measured by direct observation under the biomicroscope. The expert counts the time elapsed from the first blink until the break-up detection. This way, the test is highly dependent on the attention and the ability to react quickly of the expert. Moreover, the expert only has one opportunity to detect the break-up and measure the time, and the analysis can not be performed by other experts. In order to overcome these limitations, the fluorescent tear film can be videotaped by a camera attached to the slit lamp. Thus, the tear film video can be analyzed several times by different experts with the possibility to pause, rewind, or watch the video frame by frame.

BUT test only examines the appearance of the first dark spot in the tear film, regardless the subsequent break-up dynamics. However, break-up properties can be related to specific aspects of the tear film that could affect dry eye symptoms [King-Smith et al., 2009]. On one hand, the break-up location can

be related to tear film stability [Montes-Mico and Alio, 2005]. In this sense, the central cornea is much more sensitive than the periphery [Chang-Ling, 1989], so this area is specially vulnerable to tear disruptions. Thus, centrally located tear film break-up should affect vision as a result of tear film changes over the pupil [Liu et al., 2006]. Furthermore, the break-up sometimes occurs in the same spatial location in different measurement areas throughout the video. This may indicate a weakened area of the tear film. On the other hand, the first break-up could appear as a small point or as a wide area and it could increase its size with time presenting different rupture patterns. In the literature, three different rupture patterns have been identified as streaks, dots, or pools (see Fig. 1.14) [Bitton and Lovasik, 1998]. The streak rupture pattern has a linear shape while a circular morphology is characterized as a dot. A disturbance of the tear film that has neither a linear nor a circular shape is characterized as a pool rupture pattern.

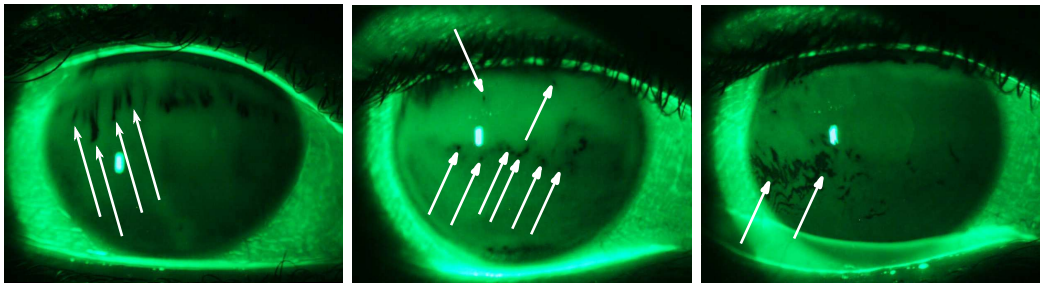


Figure 1.14: Streak, dot, and pool rupture patterns, in this order.

These patterns are related to the BUT measure, the break-up area, and its growing rate. Streaks rupture patterns present the shortest BUT while the pools and dots patterns take longer to appear. Therefore, the rupture patterns are a good indicator of tear film stability and quality, providing information about the dynamic aspects of tear rupture and morphology. Additionally, the patterns are associated with resistance to mucin contamination, so they

also may be a useful clinical indicator for identification of poor contact lenses candidates. This information is omitted in the original BUT test, but could be relevant in tear film assessment for understanding the tear film instability in DES and its relation to ocular surface symptoms [Begley et al., 2006].

The main drawback of this test is its low repeatability, mainly due to a subjective appreciation of the dark spots, the differences among the experts, and the variability of the tear film. Besides the subjectivity, the characterization by hand of break-up dynamics is a tedious and time consuming task. The automation of the break-up assessment would reduce its subjective character, allowing a more accurate evaluation of the tear film stability.

1.5 Previous Work

The automation of the BUT measure is a little explored field. To the best of our knowledge, only one methodology was proposed in the literature. This methodology [Yedidya et al., 2008, 2009] analyzes videos that contain several BUT measurements. First, the blinks are manually detected in order to delimit the sequences of interest. Then, the iris is located in the first frame of the sequence. This stage is carried out assuming the shape of the region of interest is a perfect circle. To this end, an edge map of the image immediately after a blink is created using the Canny edge detector [Canny, 1986]. In these videos, the pupil is not visible at all and the edges of the eyelids are usually stronger than the iris borders due to the fluorescein spreading. The ratio between the iris pixels and everything else is too high so it is impractical to perform a fitting to the iris directly. For this reason, three thresholded images are created over the edge image: one of the iris (I_{iris}) and the other ones (I_{low},I_{up}) of the lower and upper eyelids. Then, RANSAC algorithm [Hartley and Zisserman, 2004] is used to fit a quadratic polynomial to I_{low} and I_{up}. Then, the upper and

lower eyelids are segmented and, after that, all pixels above and below them, respectively, are discarded in I_{iris}. Finally, RANSAC is applied in order to fit a circle model to the remaining pixels.

Once the iris is well delimited in the first frame, as shown in Fig. 1.15, the extraction of the region of interest in the rest of the sequence is done in a similar way. Then, the circles representing the iris are aligned throughout the sequence, in order to make the procedure independent of slight ocular motions.

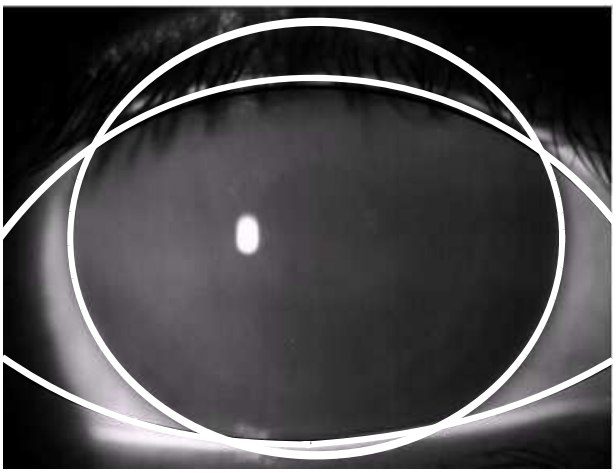


Figure 1.15: Region of interest detection from curves fitted for the eyelids and the iris

The last stage of the methodology consists in scanning the aligned video to compute the BUT measure. The break-up of the tear film is characterized by the appearance of dark spots on the surface of the eye. In order to detect the emergence of these points, the intensity evolution of each pixel is examined through the video sequence.

This approach is mainly focused on the break-up time measurement, achieving good results in this task. However, there are other features such as the rupture pattern or the size, location, and evolution of the break-up areas, that could be relevant for clinical practice. Moreover, this algorithm is not fully automatic, since the location of the BUT sequences is manually performed.

1.6 Objectives

This work presents a novel methodology for a fully automatic assessment of the tear film break-up. This evaluation includes the BUT measurement computation as well as a quantitative and qualitative analysis based on location, size, shape, and evolution of the break-up areas.

The tear film video duration is two minutes on average and contains different BUT tests. Each BUT test is a sequence delimited by consecutive blinks. The area of analysis, this is, the iris, is extracted within each frame and aligned throughout the sequence, in order to delimit the search space for break-up evaluation.

Finally, the break-up assessment is performed in each sequence. For this purpose, the BUT measurement will be computed first, as the time elapsed from the beginning of the sequence until the emergence of the first break-up in the tear film. Besides the BUT measurement, there are other break-up properties related to tear film stability, so the break-up dynamics will be characterized until the end of the measurement area. The region where the tear film breaks is related to its quality, so a local BUT analysis could provide useful additional information about the break-up location. Furthermore, the break-up could appear as a small point or a wide region, and also could evolve slowly or quickly following different rupture patterns. Therefore, the position, size, growing rate, and morphology patterns of the break-up areas will be characterized for each sequence.

The evaluation of the break-up dynamics allows a quantitative, objective analysis of tear instability, as an extension of BUT measurement, which is focused only on time. It provides additional information to the clinical practice such as the detection of weakened areas of the tear film, identification of contact lenses candidates as well as dynamic indicators of severe DES.

1.7 Outline

This thesis describes the fully automatic methodology proposed for the tear film break-up assessment, including the computation of the BUT measurement as well as the analysis of the break-up dynamics in terms of location, shape, size, and growing rate

Chapter 2 focuses on the video pre-procesing steps performed to extract data of interest for the break-up assessment. It comprises the location of BUT sequences and the extraction of the area for break-up evaluation within each frame throughout the videos.

Chapter 3 is devoted to present the break-up assessment for each sequence, as well as a comparison between the different BUT sequences. This analysis includes the BUT measurement and the characterization of the break-up areas based on the location, size, shape, and evolution.

Chapter 4 shows the experiments performed to validate the methodology and the obtained results.

Finally, Chapter 5 provides a brief overview of some concluding remarks and future directions.

Chapter 2

Sequences and Regions of Interest in the Tear Film Video

To perform the BUT test, sodium fluorescein is instilled into the eye, and the fluorescent tear film is videotaped by a camera attached to the slit-lamp biomicroscope. The patient is instructed to blink three times naturally, without squeezing, in order to distribute the fluorescein over the cornea, and then, he/she maintains the eye open as long as possible. The patient repeats this procedure several times in order to have different examples of tear film dynamics. This way, each tear film video has a duration of several minutes and contains different BUT sequences. Each BUT sequence consists of a set of frames delimited by blinks where the BUT test can be performed, hereinafter, Sequence Of Interest (SOI). Thus, after the initial full blink which marks the beginning of the SOI, the fluorescein spreads uniformly over the stable tear film. As time passes by, the tear film loses stability, and dark areas appear related to the tear film break-up. These rupture zones evolve until the final blink, which marks the end of the SOI. The duration of a SOI can vary between 5 until 35 seconds. Figure 2.1 shows representative frames in a tear film video with different SOIs.

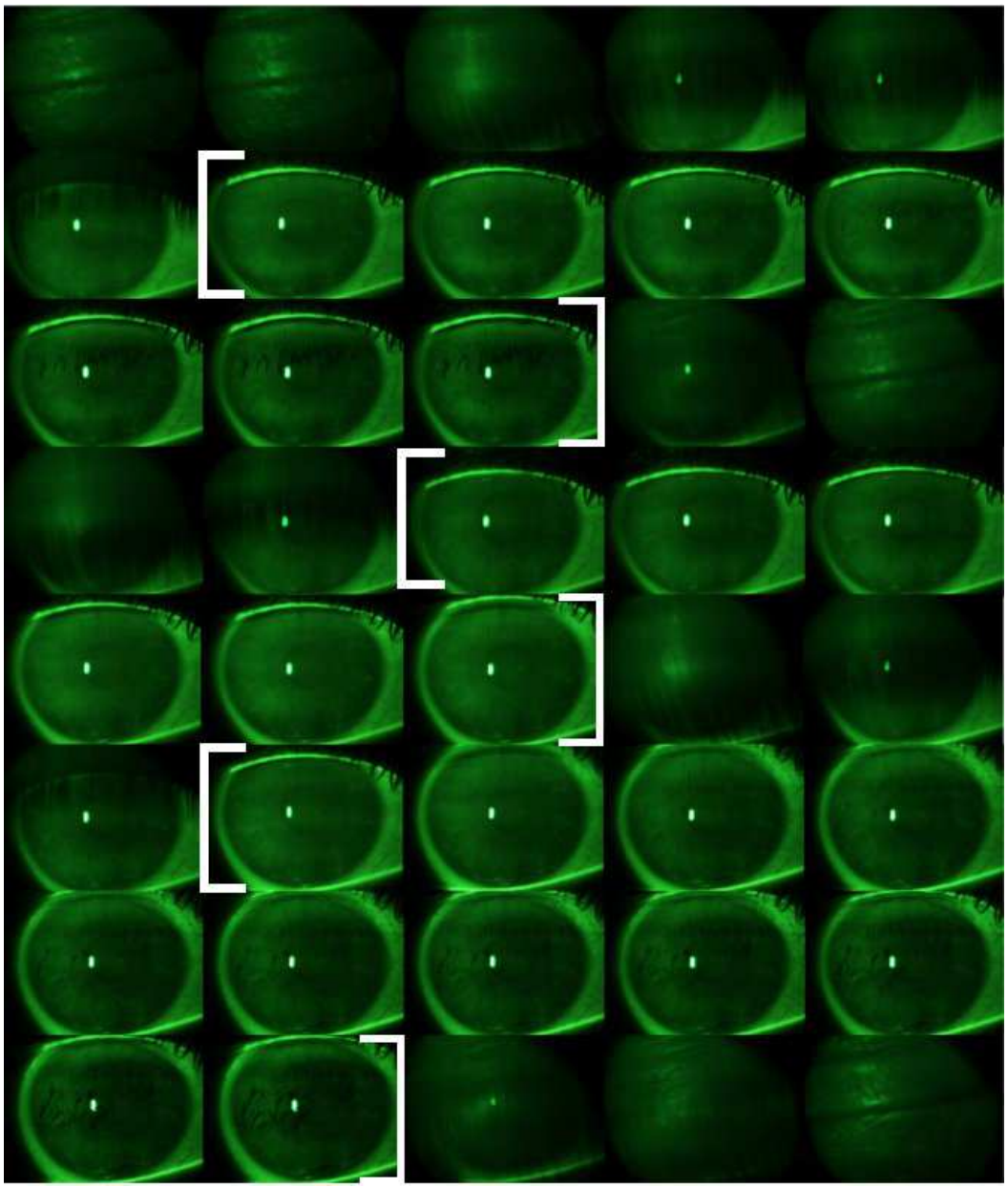


Figure 2.1: Several frames extracted at different points throughout a tear film video which contains three SOIs. White brackets mark the beginning and the end of each SOI.

The Region Of Interest (ROI) in each SOI frame corresponds to the visible part of the iris which may vary slightly throughout the sequence depending on the eye aperture and the appearance of shadows due to outer parts of the eye like eyelids or eyelashes. Figure 2.2 shows the ROI in several representative SOI frames.

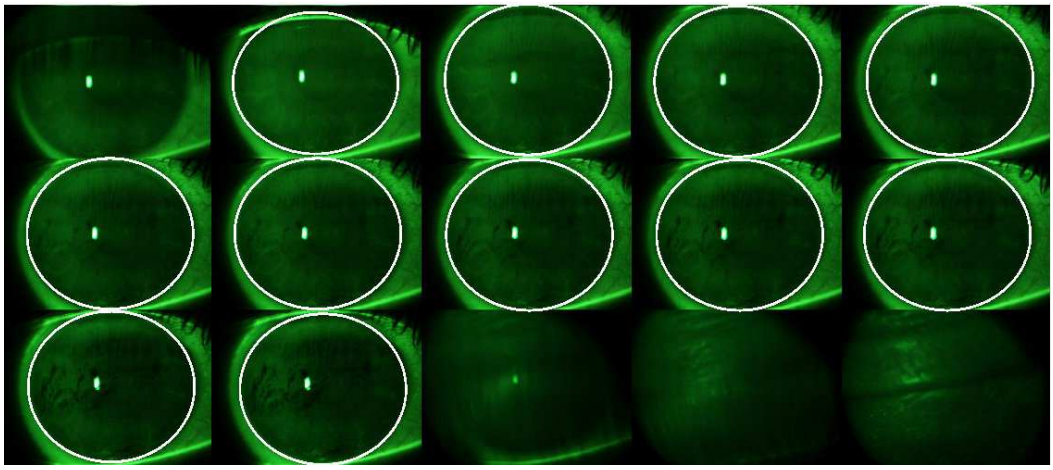


Figure 2.2: Several frames extracted at different points throughout a SOI, where tear film evolution can be observed. After the blink at the beginning of the SOI, the fluorescein is uniformly spread over the tear film, and then, as times passes by, break-up appears, increasing and becoming darker throughout the sequence until the final blink. The white ellipses mark the ROI within each SOI frame.

In order to perform the break-up assessment, the videos are preprocessed to extract the SOIs in the sequence and after that, the ROI in each frame.

The original tear film frames are images in the RGB color model with intensity values of 0 to 255. Figure 2.3 shows an example of the three components in an original frame. In this kind of images, the green component contains most of the information because the fluorescein instilled is also green.Thus, red and blue components are discarded and only the green component is used from now on. This allows to keep the necessary information in an unique component, simplifying the computation.

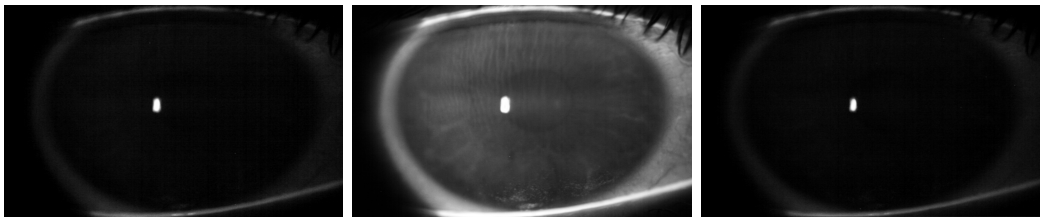


Figure 2.3: RGB components of a tear film original frame. From left to right, red, blue, and green components. Green component contains most of the information, with mean intensity of 75.39, whereas the red and blue component have mean intensities of 23.44 and 22.46, respectively.

2.1 Selection of Sequences of Interest

The different SOIs forming the tear film videos are separated by blinks which delimit their start and end frames. These blinks correspond to transitions from closed to open eye and viceversa. If the eye is open, the bright part corresponding to the sclera occupies a significant percentage of the frame, as shown in Fig. 2.4 (left). Meanwhile, if the eye is closed, the eyelid takes up the entire frame, and it presents a darker tonality, as shown in Fig. 2.4 (right).

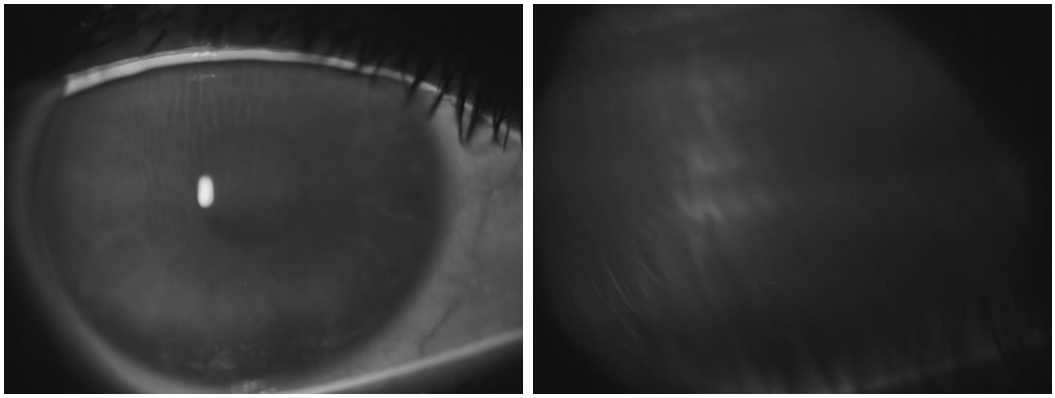


Figure 2.4: Frames during a blink in a tear film video. Left: Open eye. Right: closed eye.

Thus, a way to characterize the eye aperture is the use of a measurement which represents the frame tonality. The mean value of intensities $\overline{I_k}$ represents the average tonality of a frame and can be used to distinguish between open

and closed eyes. It is computed using the following equation [Poynton, 1999]:

$$\overline{I_k} = \frac{\sum_{i,j} I_k(i,j)}{r * c} \tag{2.1}$$

where $I_k(i,j)$ corresponds to the value of the pixel located in the row i and the column j of the frame I_k of the video sequence, and r and c are the number of rows and columns, respectively.

Once a metric for identifying open and closed eyes is defined from intensity values, the detection of blinks is based on detecting intensity variations throughout the sequence. To this end, the finite differences of mean values of gray between consecutive frames are calculated using the following equation:

$$\Delta\overline{I}_{k,d} = \overline{I}_{k+d} - \overline{I}_k \tag{2.2}$$

where $\overline{I}_k$ is the mean value of gray for the frame I_k, and d is the distance between the frames where the difference is computed. This way, $\Delta\overline{I}_{k,1}$ represents the differences between consecutive frames whereas $\Delta\overline{I}_{k,2}$ stands for the differences between every two frames. Since a blink can last several frames, the computation with a d larger than 1 is required. Figure 2.5 represents the differences between consecutive frames in a tear film video. On these differences, the peaks represent the blinks whereas the flat areas are related to the time while the patient keeps the eye open.

This way, blinks are related to peaks which represent sudden changes in the mean intensity. In order to identify these peaks, a threshold t_Δ is defined. For this purpose, two options can be considered. A fixed threshold can be used as first approach. However, each tear film video presents different lighting conditions, and the amount of fluorescein instilled varies, so intensity mean values change considerably in different videos. Thus, given the variability in illumination conditions, an adaptive threshold can be computed according to

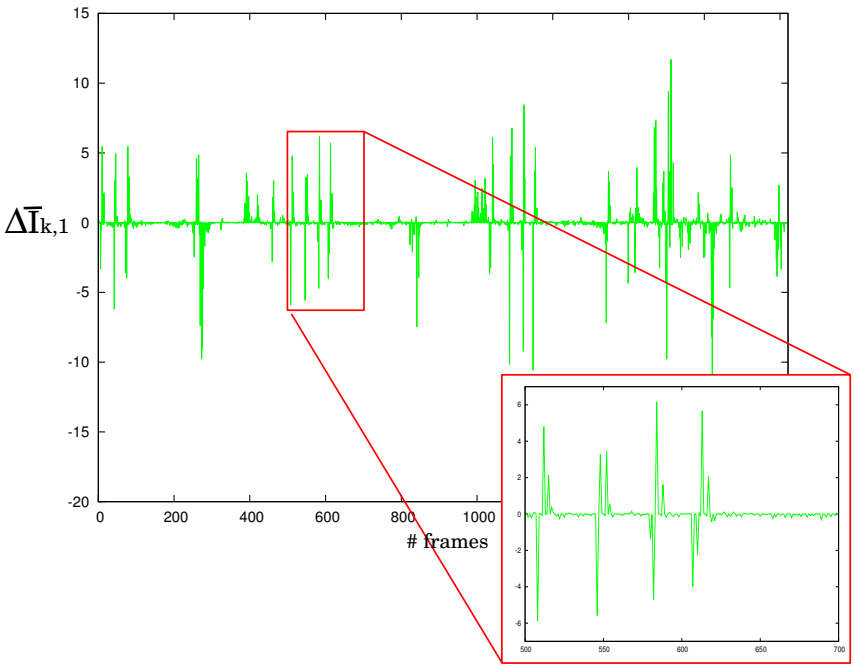


Figure 2.5: Symmetric finite differences of mean values of gray between consecutive frames. The outside graph contains a zoom of the region marked in red, in which the peaks related to different blinks can be appreciated.

the differences obtained for each video, using the following equation:

$$t_{\Delta} = \frac{\sum_{k=1}^{n-d} \Delta\overline{I}_{k,d}}{n-d} \tag{2.3}$$

where $\Delta\overline{I}_{k,d}$ are the differences of mean values of gray between consecutive frames, n is the number of frames in the tear film video, and d is the distance between the frames where the difference is computed.

The small differences related to slight movements produced while the eye remains in the same state could penalize the global calculation of t_{Δ}. In order to discard these small values, differences minor than 1 are dropped from this calculation, so Eq. 2.3 is rewritten as follows:

$$t_{\Delta} = \frac{\sum \Delta\overline{I'}_{k,d}}{|\overline{I'}_{k,d}|}, \overline{I'}_{k,d} = \{\overline{I}_{k,d} > 1, k = 1..n-d\} \tag{2.4}$$

The range $(-t_{\Delta}, t_{\Delta})$ obtained from this threshold is used to identify transitions from open to closed eye and viceversa. Thus, values outside this range are related to blinks, while values inside this range correspond to areas where the eye remains in the same state, as shown in Fig. 2.6.

On one hand, a negative peak represents the beginning of a blink, since there is a transition from a lighter frame (open eye) to a darker frame (closed eye). On the other hand, at the end of the blink there is a transition from a darker frame to a lighter frame, producing a positive peak. In some cases, there could be two consecutive differences with the same sign. This occurs when the lamp is off or when there are semi-blinks in the SOI. In this case, the lowest absolute values are removed until all pairs of consecutive blinks have opposite signs, as shown in Fig. 2.7.

Once the peaks with opposite signs are identified, the SOIs are extracted as those intervals starting with a positive difference and ending with a negative difference. Furthermore, the SOIs should exceed a length t_L to ensure that

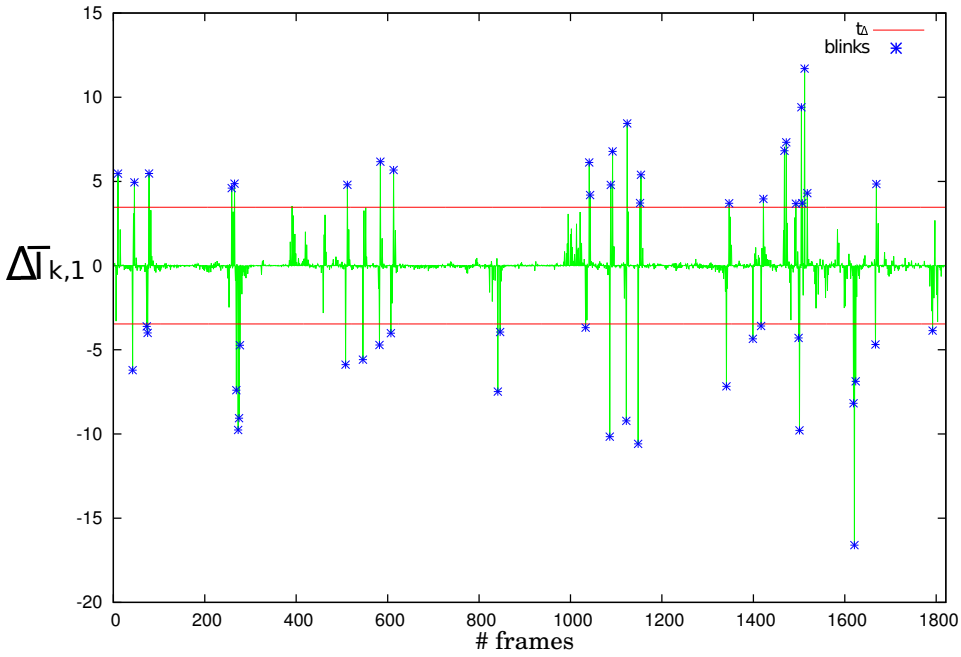


Figure 2.6: Blink extraction from symmetric finite differences of mean values of gray. Peaks correspond to blinks while flat areas are related to frames where the eye remains open or closed.

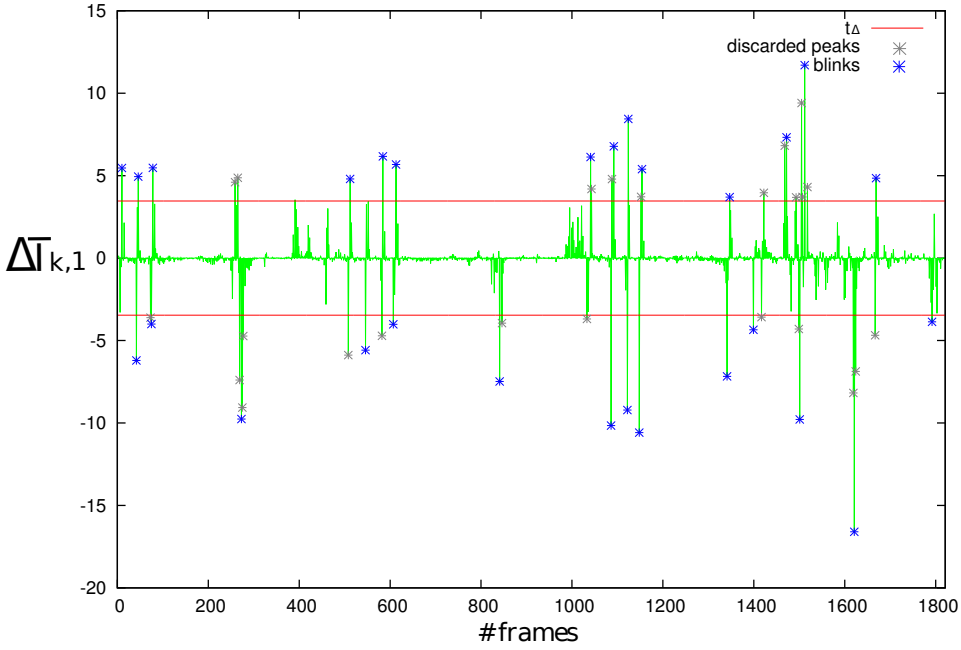


Figure 2.7: Discarding consecutive peaks with the same sign until all pairs of consecutive blinks have opposite signs.

they have a minimum of frames for break-up assessment. Figure 2.8 shows the SOIs detected after discarding the invalid intervals.

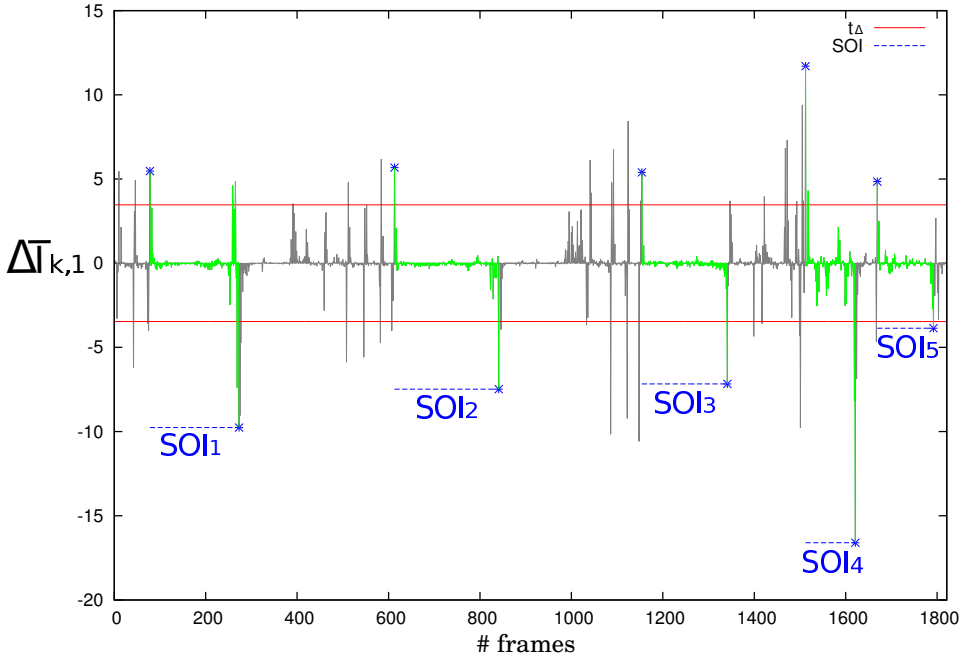


Figure 2.8: Automatic detection of SOIs as the intervals between peaks with opposite signs which exceed a length t_L.

Sometimes the blink occurs gradually and the differences between successive frames may not be enough to detect it. In order to detect all the blinks, differences between non-consecutive frames, separated by a distance of f frames, are also considered. Thus, $\Delta\overline{I}_k^f$ is computed as the sum of finite differences between the frame i and the frame $i+d$, where d gets values from 1 to f frames:

$$\Delta\overline{I}_k^f = \sum_{l=1}^{f} \Delta\overline{I}_{k,l} \tag{2.5}$$

This sum emphasizes the intensity changes produced during a blink. Then, t_Δ is obtained as follows:

$$t_\Delta = \frac{\sum \Delta\overline{I'}_k^f}{|\overline{I'}_k^f|}, \overline{I'}_k^f = \{\overline{I}_k^f > 1, k = 1..(n-d)\} \tag{2.6}$$

Figure 2.9 (top) shows the detection of a SOI using the finite differences with f set to 1. In this case, the delimited SOI includes a few frames belonging to a blink not detected with this approach. However, in Fig. 2.9 (bottom), the detection is performed from the sum of differences with f set to 2, which allows to detect this blink and delimit accurately the SOI.

2.1.1 Reference Frames in SOI

At the beginning of each SOI there are several frames after the initial blink in which the patient is still opening the eye. In the same way, the blink which delimit the end of the SOI takes some frames while the eye is closing. For this reason, an adjustment of the SOI is performed to discard these transition frames. It consists in detecting the first frame in which the eye is fully open, defined as the Start Reference Frame (SRF) and the last frame before the eye begins to close, defined as the End Reference Frame (ERF). Since the blink duration is not fixed, as can be seen in Fig. 2.10, adaptive SRF and ERF frames are selected for each SOI. To this end, the sum of finite differences $\overline{I}_k^f$, computed in the previous step, is used.

The curve starts its evolution with a positive peak related to the eye fully closed. The value decreases while the eye is opening, until it reaches a local minimum value that corresponds to the eye fully open. The frame related to the first local minimum of this curve is selected as the SRF. This is the point where the analysis of the evolution of fluorescein in the tear film begins. On the other hand, when the eye starts to close at the end of the SOI, the negative differences $\overline{I}_k^f$ increase until they reach a peak related to the eye fully closed. The frame related to the local maximum value of the decrease previous to this negative peak is selected as the ERF. This is the final point for the break-up analysis.

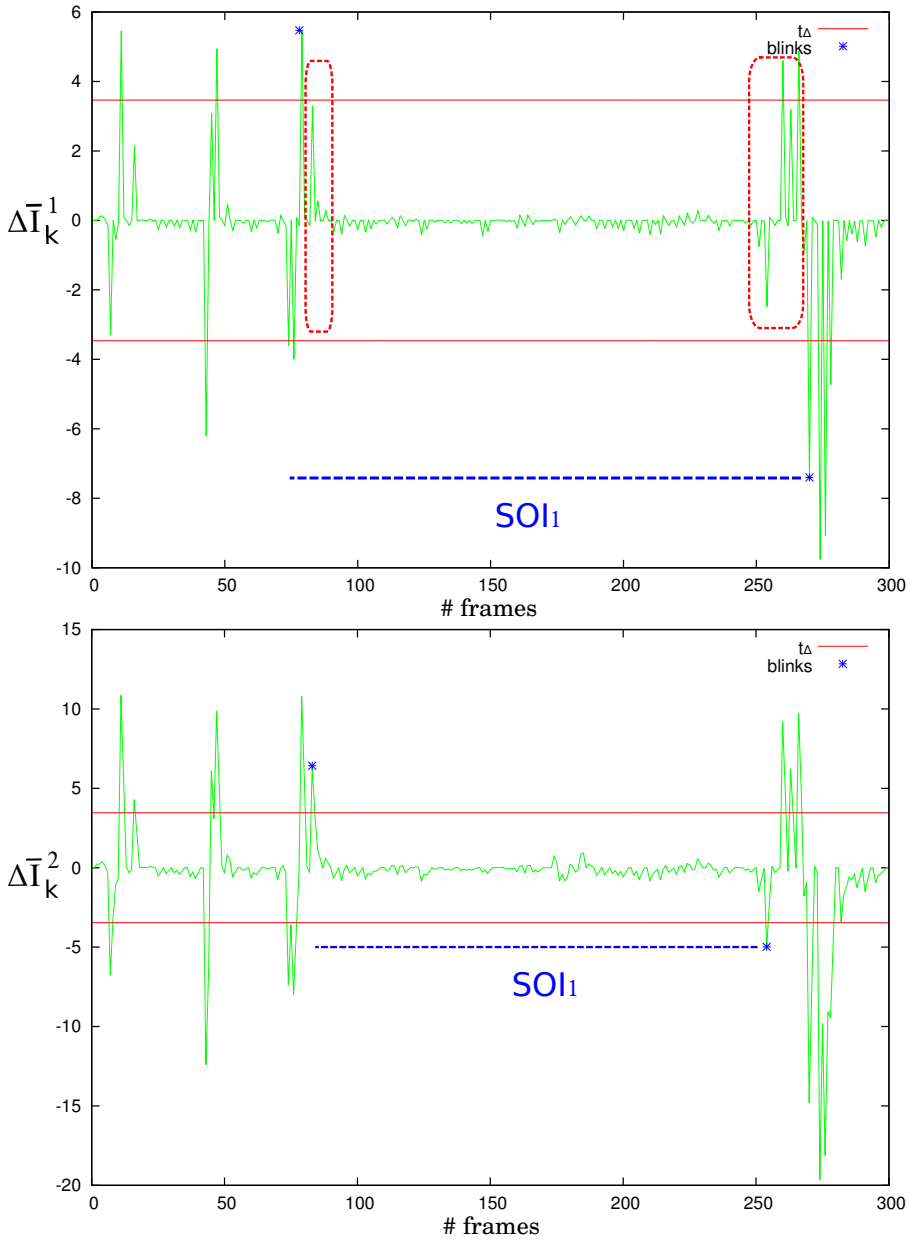


Figure 2.9: Detection of a SOI using symmetric finite differences of mean values of gray. Top: differences between consecutive frames ($f = 1$). Bottom: sum of the differences with $f = 2$.

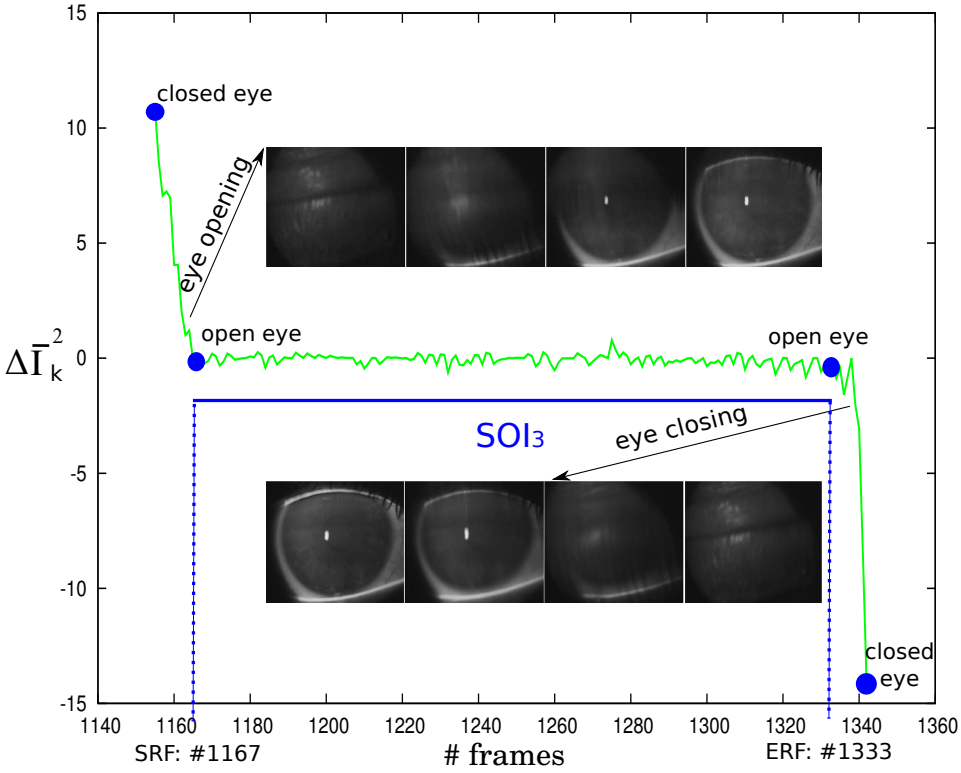


Figure 2.10: Evolution of the $\overline{I}_k^f$ during the stabilization after a blink. The negative peak correspond to the closure of the eye while the positive peak correspond to the opening of the eye.

2.2 Extraction of the Region of Interest

Once the SOIs are located, the next stage is the detection of the region of interest (ROI) within each frame to discard frame regions with no information. The main steps of this procedure are shown in Fig. 2.11. First, the position and size of the iris are identified, since it corresponds to the region where the break-up is evaluated. Since the eye does not remain fixed in the same position throughout the video, the next step consists of aligning the ROI in each frame. This way, the methodology is independent of slight motions of the ROI. Moreover, video frames include outer parts of the eye, like eyelids or eyelashes, which not contain relevant information and could mislead the results. Thus, in the last step, a ROI adjustment is performed according to the eye features, in order to discard these outer regions.

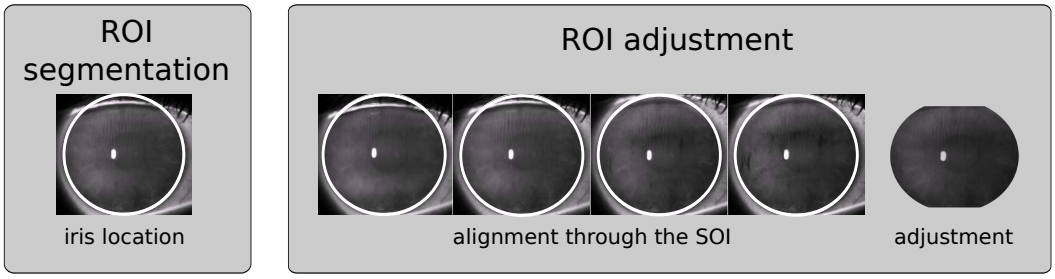


Figure 2.11: Steps for extracting the ROI in each SOI.

2.2.1 ROI segmentation

Assuming the iris has an approximately circular shape (see Fig. 2.12, a), its initial segmentation is carried out by correlation of an edge image with a set of masks formed by circumferences with different radii covering typical eye sizes. A mask of radio R centered at the midpoint of the frame (c_i, c_j) is defined as

follows:

$$M_R(i,j) = \begin{cases} 1 & if \quad (i-c_i)^2 + (j-c_j)^2 = R^2 \\ 0 & otherwise \end{cases} \tag{2.7}$$

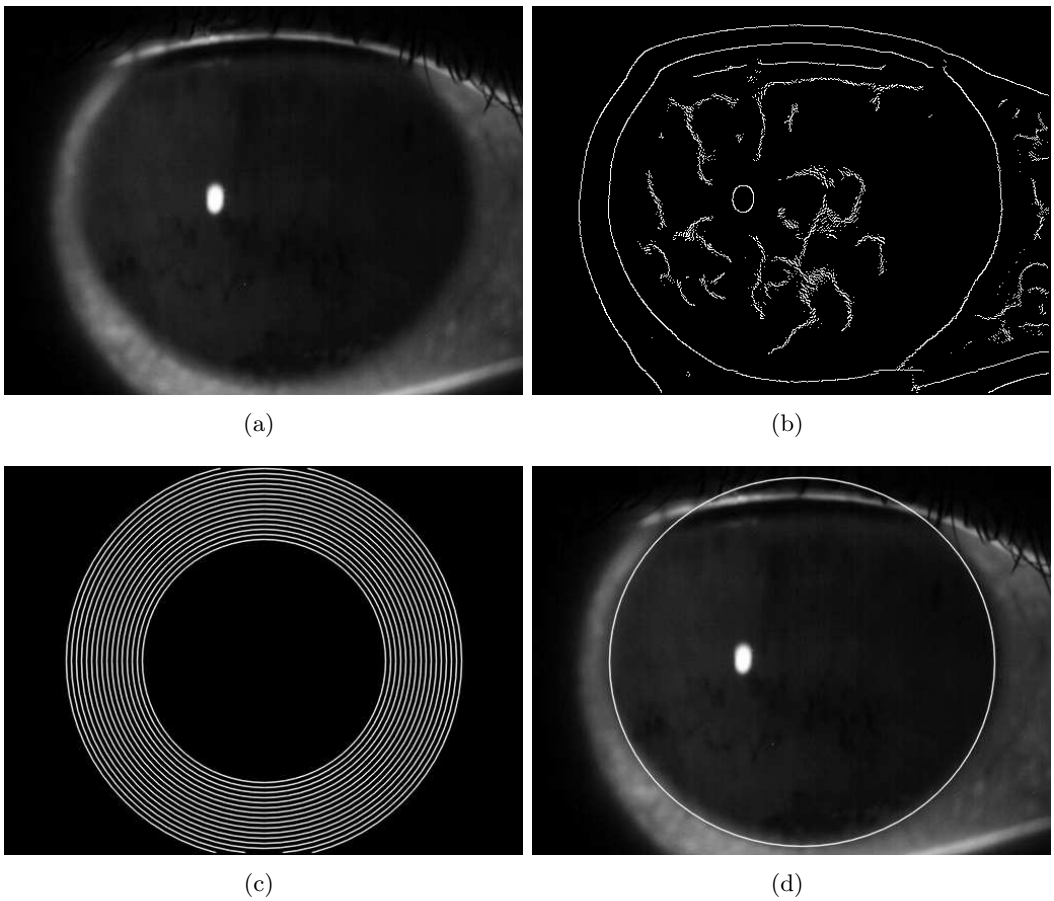
The Canny edge detector [Canny, 1986] is applied to the green component (see Fig. 2.12, b), and then, the normalized cross correlation is computed in the frequency domain between the edge image and the set of masks (Fig. 2.12 (c)):

$$NCC_R(x,y) = \frac{\sum_{i',j'}(M_R(i',j') \cdot E(i+i',j+j'))}{\sqrt{\sum_{i',j'}(M_R(i',j')^2 \cdot \sum_{i',j'} E(i+i',j+j')^2}} \tag{2.8}$$

where M_R represents the mask image with dimensions $w \times h$ formed by a circumference of radius R, E is the edge image corresponding to the tear film frame in which we expect to find the best match, and NCC_R stores the match for each location result of sliding M_R over E. The sliding is done one pixel at time (left to right, up to down), with $i' = 0..w-1$ and $j' = 0..h-1$. The correlation for each mask is obtained from the maximum value in NCC_R, since it corresponds to the best match. Therefore, the radii and position of the mask with the highest matching value delimit the iris, as Fig. 2.12 (d) shows.

The natural structure of the eye contains the sclera, a white fibrous membrane that forms a circular region surrounding the iris, which also appears in the edge image, as Fig. 2.13 (a) shows. Sometimes, the masks match this region instead of the iris, as can be seen in Fig. 2.13 (c). In order to avoid these mismatches, the edge pixels corresponding to the sclera area are discarded. To this end, the orientation of the Canny edge image is considered. Thus, instead of correlating the original edge image E, an edge image E' which only contains the outgoing edges in relation to the frame center (c_i, c_j) is obtained as follows:

$$E'(i,j) = \begin{cases} 1 & if \quad E(i,j) = 1 \wedge j \geq c_j \wedge \theta(i,j)\epsilon(\frac{\pi}{2}, \frac{3\pi}{2}) \\ 1 & if \quad E(i,j) = 1 \wedge j < c_j \wedge \theta(i,j)\epsilon(\frac{3\pi}{2}, \frac{\pi}{2}) \\ 0 & otherwise \end{cases} \tag{2.9}$$



(a) (b)

(c) (d)

Figure 2.12: ROI extraction by correlation between a set of circumferences and the edge image. (a) Green component from the original image. (b) Edge image. (c) Set of masks for correlating to the edge image. (d) ROI delimited from the best match.

where $E(i, j)$ is the original edge image and θ is the orientation of the edge pixels obtained. Appendix B describes the adapted Canny edge detector. Figure 2.13 (b) shows the edge image after discard the incoming edge pixels, and Fig. 2.13 (d) shows the correct match obtained by including this improvement.

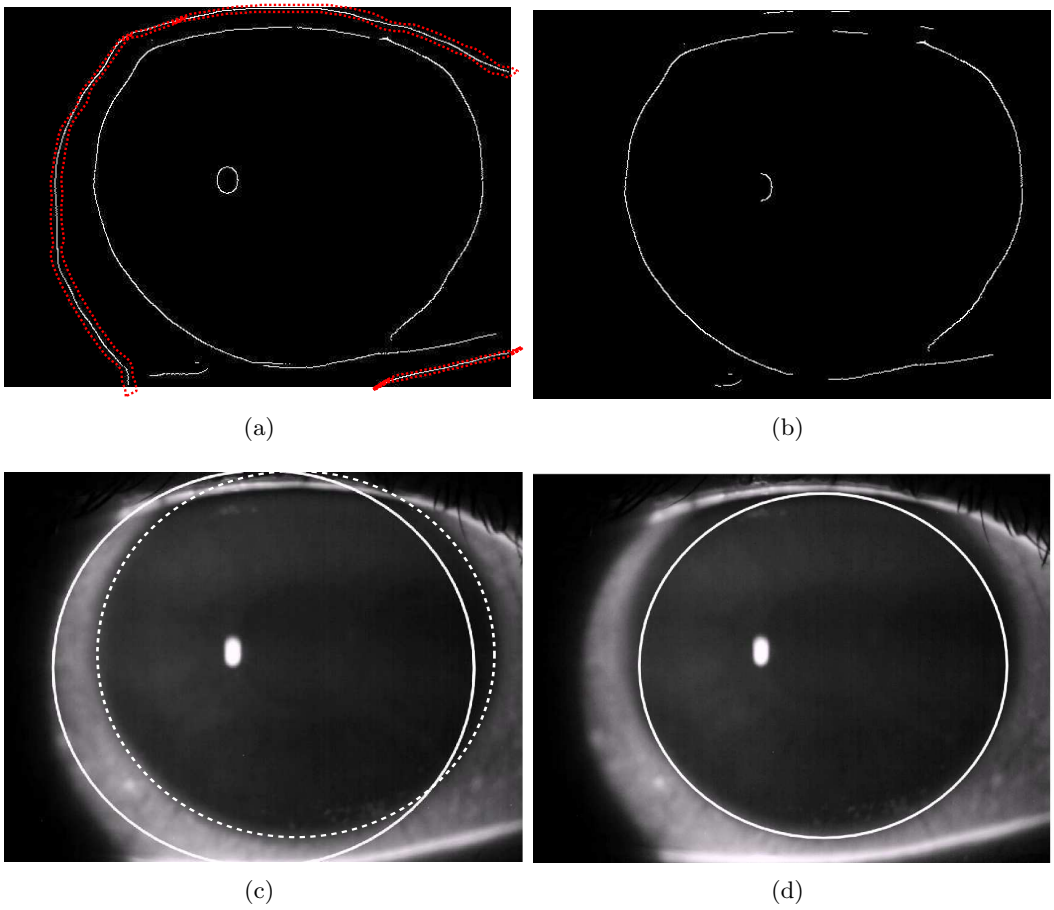


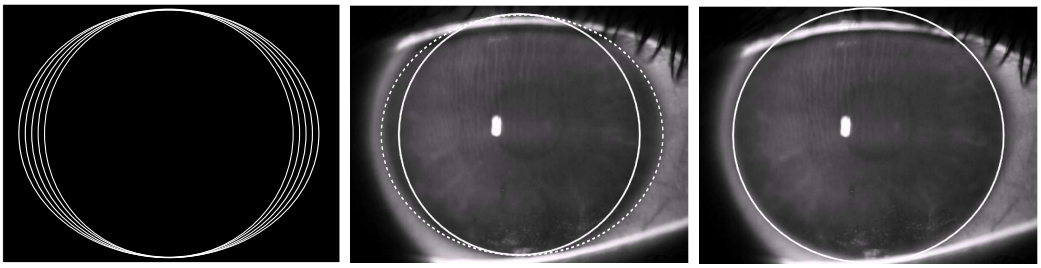
Figure 2.13: ROI extraction by correlation to the edge image. (a) Original edge image which contains edge pixels related to the eye sclera, marked by dotted red lines. (b) Edge image discarding incoming pixels. (c) Best match computed from (a). The dotted line represents the correct ROI while the continuous line is the ROI extracted by the algorithm. (d) Best match computed from (b) where only the outgoing edges are considered.

On the other hand, the size of the eye is estimated assuming that the iris has a circular shape. In some cases, this assumption would miss a small part of the ROI, since the iris is slightly oval, as shown Fig. 2.14 (b). In order to cover

the maximum possible ROI, elliptical masks are also considered, in addition to the circumferences. Thus, each radius considered in the set of circumferences is slightly extended in the horizontal axis to create another subset of elliptical masks, as shown Fig. 2.14 (a). The elliptical masks centered in the midpoint (c_i, c_j) are built as follows:

$$M'_{R1,R2}(i,j) = \begin{cases} 1 & if \ \frac{(i-c_i)^2}{R_1^2} + \frac{(j-c_j)^2}{R_2^2} = 1 \\ 0 & otherwise \end{cases} \tag{2.10}$$

where R_1 and R_2 are the horizontal and vertical semi axis. This improvement allows to extract a larger area of relevant information, as Fig. 2.14 (c) shows.



(a) (b) (c)

Figure 2.14: ROI extraction. (a) Set of circular and elliptical masks. (b) The dotted line represents the correct ROI while the continuous line is the ROI extracted using a circumference as mask. (c) ROI extracted using an elliptical mask.

The aim of this step is extracting the ROI within each frame in the SOI. To this end, the described method can be used for adjusting the iris size in the whole SOI. However, the correlation between the set of masks and each frame through the SOI has very high computational cost. The complexity of this operation is $O(N_E \cdot N_C \cdot n)$, where N_E corresponds to the size of the edge image E, that is $(i \times j)$, N_C is the number of times the correlation is computed, that is, $(i' \times j')$, and n is the number of frames in the set of masks. Since the size of the eye does not change through the video, computing the correlation in only one frame would be enough to get the ROI. However, considering that

there may be slight variations due to noise or in the degree of eye aperture, three sample frames are considered for extracting the ROI robustly without affecting performance. Thus, the frames placed at 25%, 50%, and 75% of each SOI are selected, and the correlation process previously described is carried out within these three frames. Then, the maximum correlation in the frequency domain among the three edge images and the different masks is analyzed to determine the best fit. The largest value of these correlations corresponds to the optimal radii, which correspond to the size of the eye along the SOI.

2.2.2 ROI adjustment

The eyes are never in a state of complete rest, and even when they are fixated on one point, random jitter movements are produced in order to satisfy the demand of the photoreceptors for non-constant stimulus [Day and Brown, 2001; Carlson, 2005]. In order to discard these slight motions throughout the video, the ROI is registered throughout each SOI. To this end, the elliptical mask selected in the first step is correlated to the edge image of the SRF. Then, the SRF is cropped around a bounding box which is defined as a rectangle that contains the ellipse and is centered at the pixel with highest correlation. The next frame in the SOI is correlated to this bounding box, and a new bounding box is obtained. This procedure is repeated until the end of the SOI, correlating each edge frame to the previous bounding box, and extracting the bounding box for the current frame from the higher correlation. Figure 2.15 shows some examples of bounding boxes defined over different frames of the SOI.

After the registration step, the elliptical mask is applied since the ROI is always located in the same position. Therefore, the methodology is independent of slight motions of the ROI, since the edge alignment between consecutive frames produces a good enough match. Figure 2.16 shows an example of ROI registration.

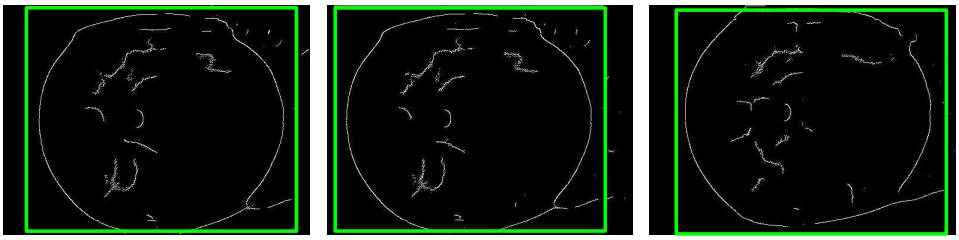


Figure 2.15: Bounding boxes defined for aligning the ROI throughout the SOI. Each bounding box, represented as a green rectangle, is centered at the ellipse which represents the iris.

In some cases, the eye is not fully open and the ROI contains outer parts like eyelids or eyelashes. These regions do not contain relevant information for the analysis and could mislead the results. For these reasons, a further adjustment is performed in each SOI frame, as shown Fig. 2.17. It consists in cropping the top and the bottom of the circular ROI, and reducing slightly the radius to get rid of noise at the boundaries of the iris.

A simple approach can be based in a fixed adjustment where the cropping parameters are obtained from the eye radii previously computed. However, each eye has a different shape and aperture degree, so an adaptive adjustment is also proposed taking into account the features of each eye. Therefore, an upper and a lower limit are calculated to crop the ROI at the top and at the bottom ends, respectively. This limits are calculated over the SRF since it is the first frame in the SOI with the eye fully open. The number of edges is computed in each row i of the SRF, as Eq. 2.11 shows. Then the upper limit is selected as the furthest row in the upper half of the image that accumulates more edge points than a variable threshold (Eq. 2.12). In the same way, the lower limit is the closest row in the lower half of the image that accumulates more edge points than a variable threshold (Eq. 2.13).

$$Acc(i) = \sum_{j=i}^{c} E(i,j) \tag{2.11}$$

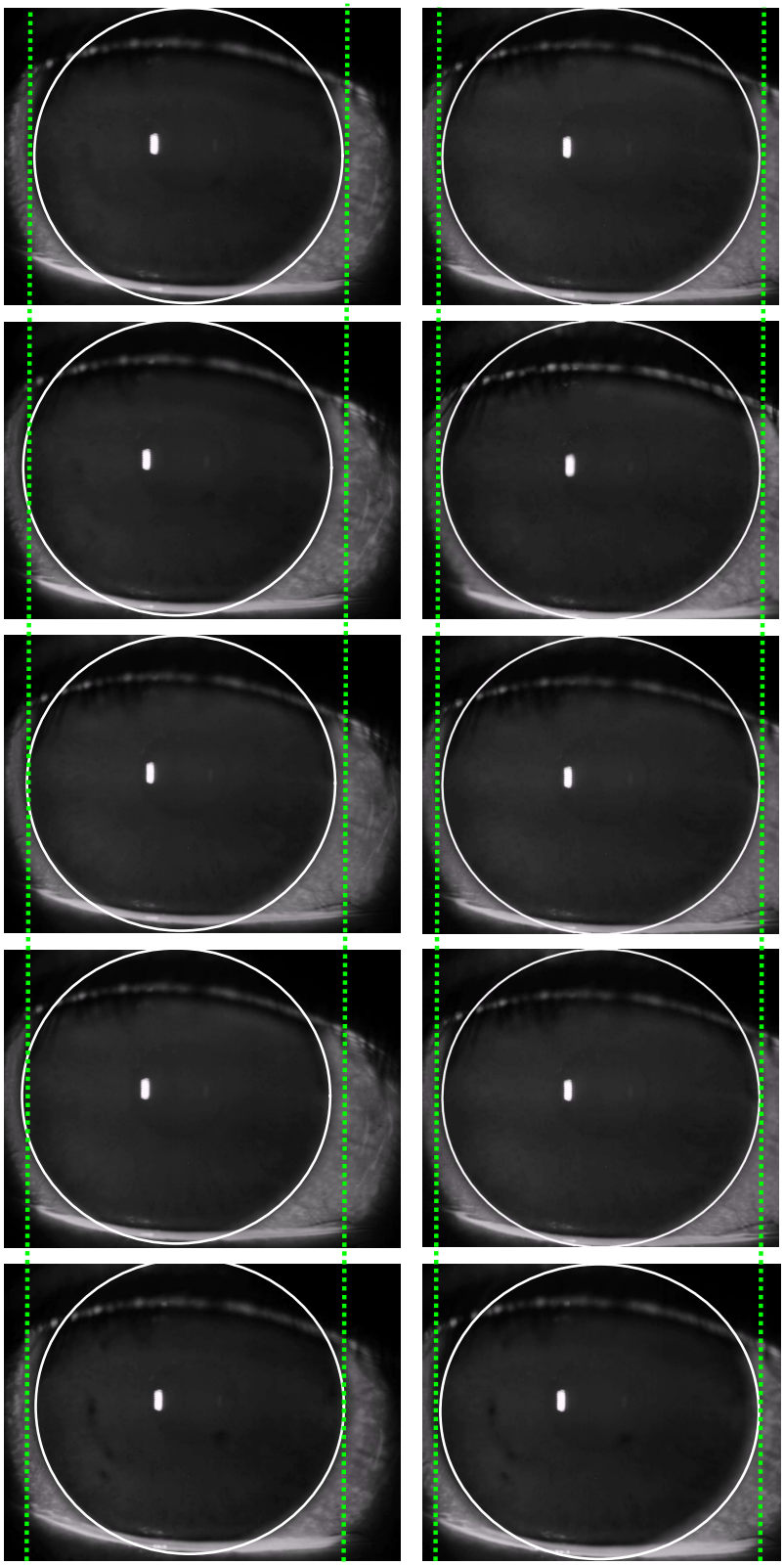


Figure 2.16: ROI registration throughout the SOI. First column represents the ROIs over a set of frames in the SOI. Due to slight motions of the eye, the ROI are not always in the same frame location. Second columns represents the ROIs over the registered frames.

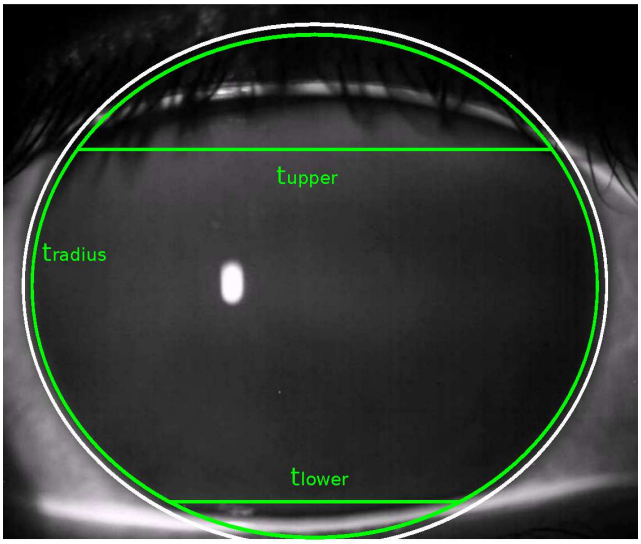


Figure 2.17: Adjustment of the ROI to discard irrelevant information. The top and the bottom are cropped, and the radii are slightly reduced.

$$l_{upper} = \arg\max_i \{Acc(i) > t_{upper}\} \tag{2.12}$$

$$l_{lower} = \arg\min_i \{Acc(i) > t_{lower}\} \tag{2.13}$$

The thresholds are computed as a percentage of the maximum number of edge points found in each half of the image. Given the nature of the eye, the upper eyelid and eyelashes usually invade more ROI than the lower ones, so the parameters α and β are used as weights for the upper and lower threshold:

$$t_{upper} = \alpha \cdot max\{Acc(i), 0 \leq i < r/2\} \tag{2.14}$$

$$t_{lower} = \beta \cdot max\{Acc(i), rows/2 \leq i < r\} \tag{2.15}$$

Furthermore, the radii which define the fitted ellipse are slightly reduced by applying the weight γ, as follows:

$$R_1' = R_1 \cdot \gamma \qquad R_2' = R_2 \cdot \gamma, \quad \gamma \in (0, 1] \tag{2.16}$$

Figure 2.18 shows the adjustment of the ROI for two SOIs, one with the eye slightly closed and the other with the eye fully open. In the first case, the eye is a bit closed and the adaptive adjustment discards these outer parts. However, in the second case the eye is fully open, so the adaptive adjustment covers more iris area.

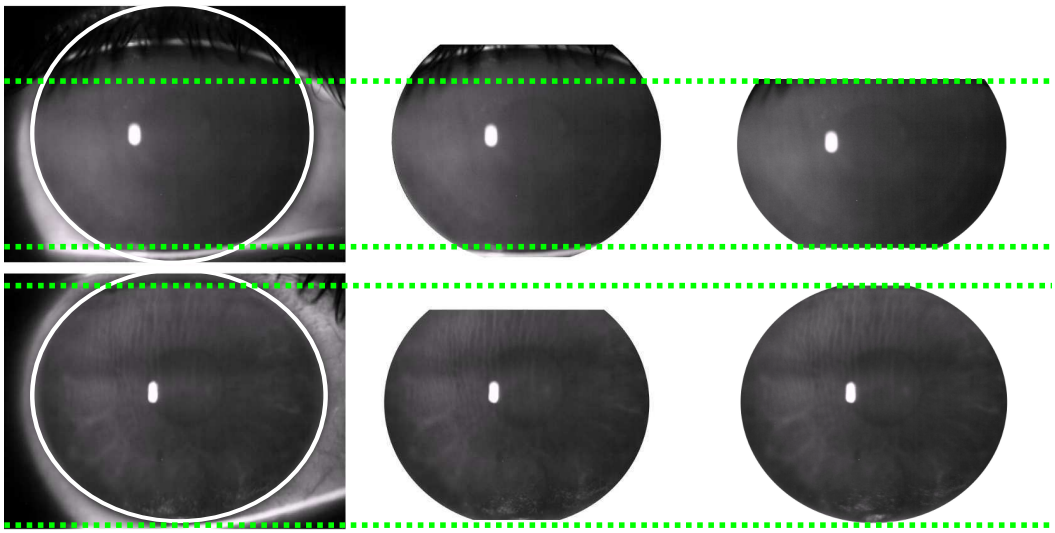


Figure 2.18: ROI adjustment. First column represents the original frames and second and third columns contain the results of applying fixed and adaptive adjustment, respectively. Dotted green lines represent the adjustment marked by the experts in the original images. In the first row, fixed adjustment includes irrelevant elements which were discarded in the adaptive adjustment. In the second row, fixed adjustment discards valid parts for the analysis, while the adaptive adjustment is similar to the region marked for the experts.

Chapter 3

Break-Up Assessment

Once the SOIs are delimited and the ROI is extracted within each frame, the key stage of the methodology is the break-up assessment. It consists of the analysis of the different SOIs to detect the break-up time as well as characterize its shape, size, location, and evolution through time. The main steps for the break-up assessment are shown in Fig. 3.1.

First, a preprocessing step is performed to normalize the tear film frames and determine a suitable threshold to segment the break-up areas. For this purpose, the contrast and illumination variability are reduced by a correction process. Then, a break-up threshold t_b is computed for each SOI to determine if a pixel corresponds to a break-up area or not. After this, the break-up is analyzed in different ways. On one hand, the BUT measurement is computed as the time elapsed from the beginning of the SOI until the emergence of dark spots. Once the BUT is measured, the break-up areas are analyzed from this frame until the end of the SOI. This evaluation includes a rupture pattern classification as well as a characterization of dynamics in terms of size and growing rate. Moreover, a local analysis in basis of Cornea and Contact Lens Research Unit (CCLRU) standards [Terry et al., 1993a,b] is also performed.

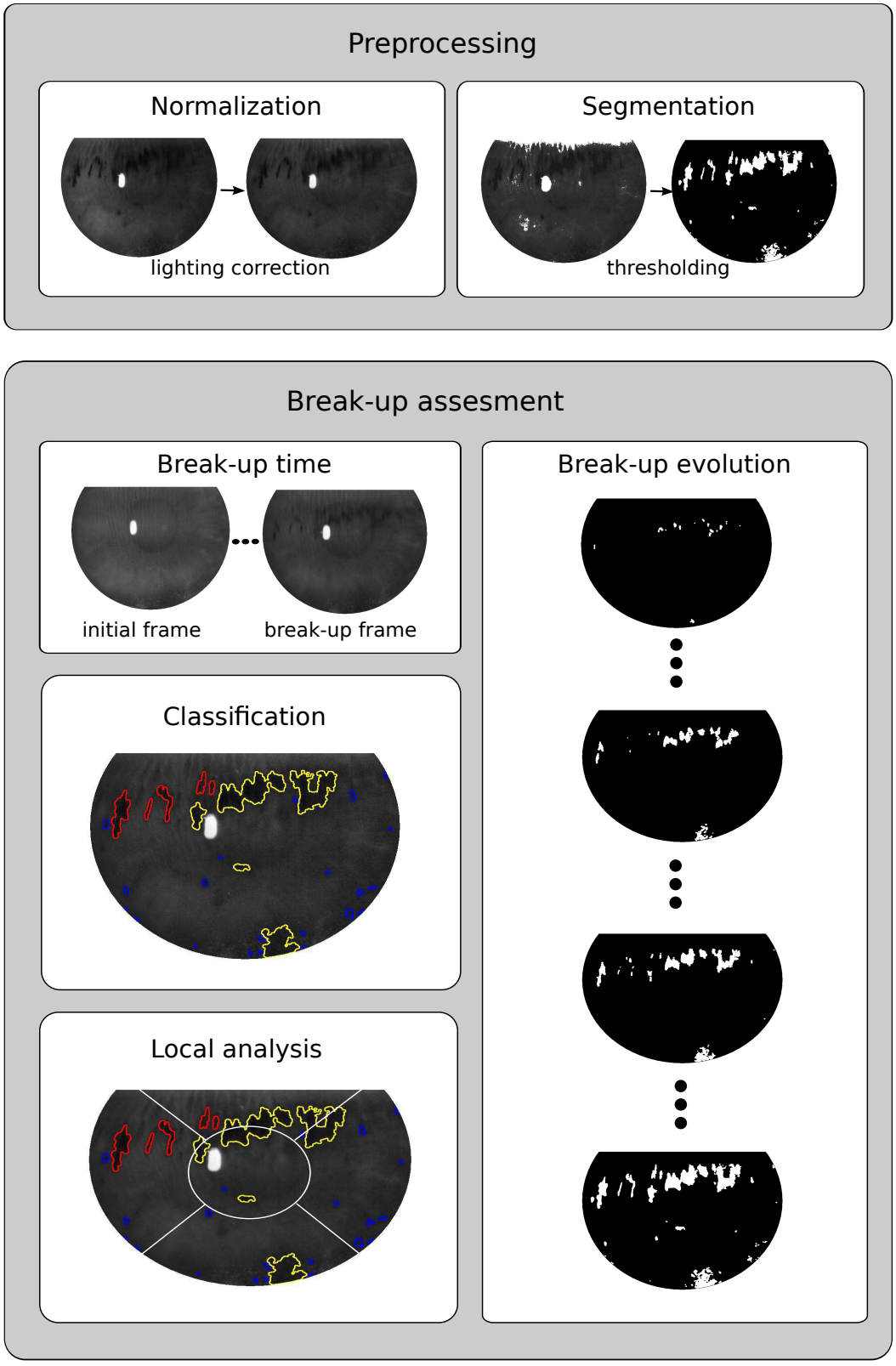


Figure 3.1: Steps for break-up assessment in each SOI

3.1 Lightness and Contrast Normalization

The video acquisition processes and the spherical surface of the eye cause luminosity and contrast heterogeneity within the tear film frames [Dror et al., 2001]. Figure 3.2 shows two examples of frames where the right part of the ROI is slightly darker than the left part. This problem could affect the break-up characterization, mistaking poorly illuminated areas with real break-up areas.

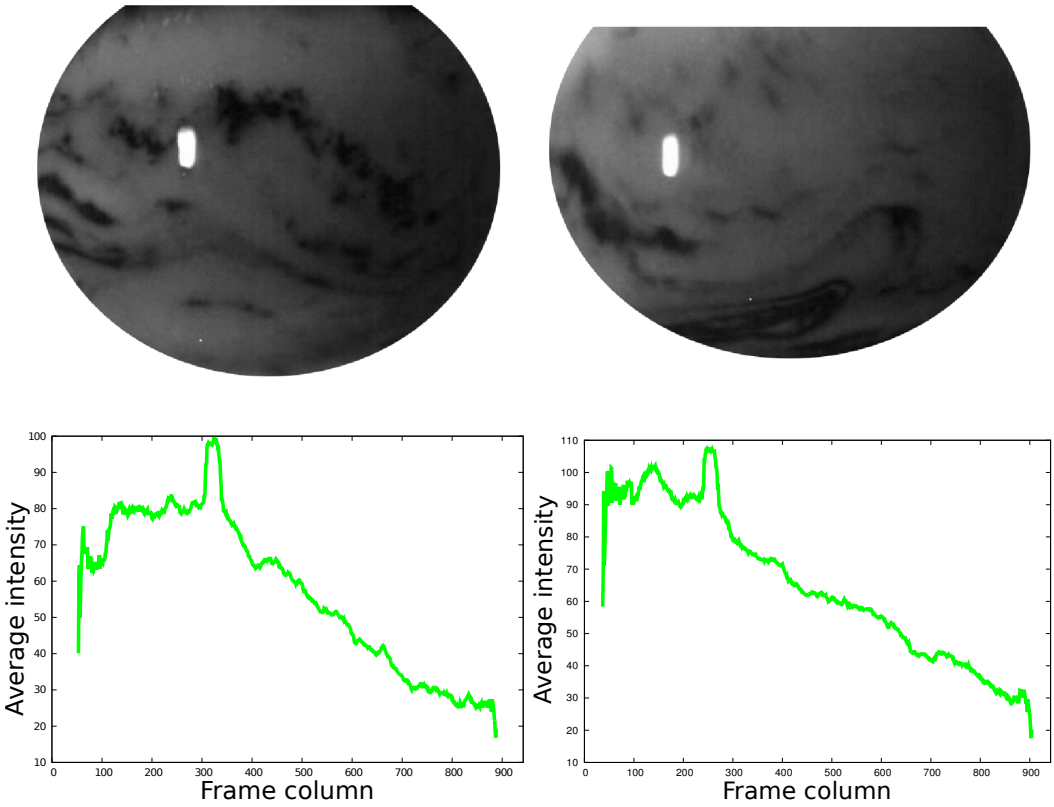


Figure 3.2: Top: Original ROI frames showing luminosity and contrast heterogeneity. Bottom: average intensity by frame columns. The left side of the image has higher intensity values than the right side.

In order to overcome this problem, a lighting correction is performed over the ROI of the SOI frames. This process consists in normalizing the lightness and contrast variability in images, based on estimating both features in small

background areas, spreading to the whole image, and then removing from it [Foracchia et al., 2005]. The method assumes the following model of the image:

$$I = f(I^\circ) = f(I_b^\circ + I_f^\circ) \tag{3.1}$$

where I° is the original image, I_b° is the original background, that is, the ideal background which represents the ROI uniformly stained by fluorescein, free of the break-up areas which are modeled in the original foreground, I_f°, and f represents the transformation made at the acquisition process which produces the lightness and contrast deformation.

The original background image follows a normal distribution with mean the ideal uniform lightness, μ_b, and standard deviation, σ_b, the natural variation in the dye's spread, $(I_b^\circ \sim \mathcal{N}(\mu_b, \sigma_b))$.

A captured image is modeled as follows:

$$\begin{aligned} I(i,j) &= f(I^\circ(i,j)) = C(i,j)I^\circ(i,j) + L(i,j) \\ &= C(i,j)I_b^\circ(i,j) + C(i,j)I_f^\circ(i,j) + L(i,j) \end{aligned} \tag{3.2}$$

where $C(i,j)$ and $L(i,j)$ represent the deviation factor of contrast and lightness respectively, which can be considered as images since they are space-dependent functions.

Thus, the normalized image is computed estimating the contrast and lightness, $\hat{C}$ and $\hat{L}$, as follows:

$$\hat{I}^\circ(i,j) = \frac{I(i,j) - \hat{L}(i,j)}{\hat{C}(i,j)} \tag{3.3}$$

Taking into account only the background pixels $(i,j) \in \mathcal{B}$ where, by defi-

nition, $I_f^\circ = 0$, the equation 3.1 is as follows:

$$I(i,j) = C(i,j)I_b^\circ(i,j) + L(i,j) \ \ , (i,j) \in \mathcal{B} \tag{3.4}$$

From the previous equation and considering the statistical model of I_b°, $(I_b^\circ \sim \mathcal{N}(\mu_b, \sigma_b))$, it is deduced that $I(i,j) \sim \mathcal{N}(L(i,j), C(i,j))$ for $(i,j) \in \mathcal{B}$. Then, the normalized image is achieved through the equation 3.3, estimating $\hat{L}(i,j)$ and $\hat{C}(i,j)$ by the mean and standard deviation of the background pixels in the observed image $(I(i,j),\ (i,j) \in \mathcal{B})$.

To estimate the mean and standard deviation in the background, first, the background pixels have to be determined. In that way, the image is divided in a tessellation of squares S_i with side s, where s is selected to obey the next three premises: L and C are constant in the square, at least the 50% of pixels in the region belong to background, and the intensity of the background and foreground pixels are very different. Then, the mean $\hat{\mu}(S_{\frac{m}{s}\times\frac{n}{s}})$ and standard deviation $\hat{\sigma}(S_{\frac{m}{s}\times\frac{n}{s}})$ images are built, computing the mean $\hat{\mu}(S_i)$ and the standard deviation $\hat{\sigma}(S_i)$ for each S_i. The full images $\hat{\mu}(S_{m\times n})$ and $\hat{\sigma}(S_{m\times n})$ are computed by bicubic interpolation from the small images $\hat{\mu}(S_{\frac{m}{s}\times\frac{n}{s}})$ and $\hat{\sigma}(S_{\frac{m}{s}\times\frac{n}{s}})$, being $m \times n$ the original image resolution. Hence, the membership to the background for each pixel in the image is determined by means of the Mahalanobis distance as follows:

$$(i,j) \in \mathcal{B} \leftrightarrow \left| \frac{I(i,j) - \hat{\mu}(S_{m\times n})(i,j)}{\hat{\sigma}(S_{m\times n})(i,j)} \right| < t \tag{3.5}$$

where t is a threshold that was set to 1, while a value of 200 for the square side was selected empirically. Figure 3.3 represents an example of background image estimation.

Once the background pixels are determined, the same tessellation of squares is taken into account and the mean and standard deviation for the background

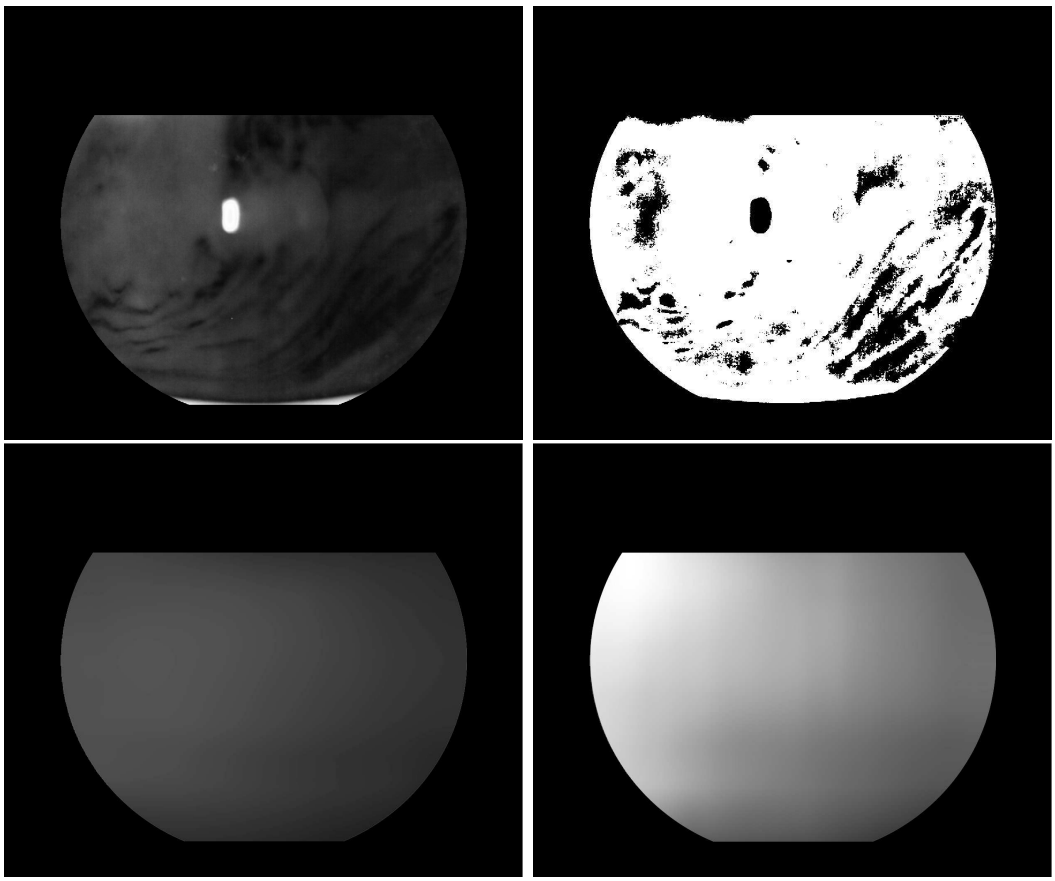


Figure 3.3: Estimation of contrast and illumination images for normalization process. From left to right, top to bottom: original frame; background image estimated from the mean and standard deviation of the squares in the tessellation; $\hat{L}(i,j)$ and $\hat{C}(i,j)$ images estimated from the mean and standard deviation of the background image.

pixels at each square are computed. Then, a bicubic interpolation is applied again to achieve $\hat{L}(i,j)$ and $\hat{C}(i,j)$. Finally, the normalized image is obtained applying the equation 3.3. Fig. 3.4 shows some examples of preprocessed frames, which present uniform luminosity.

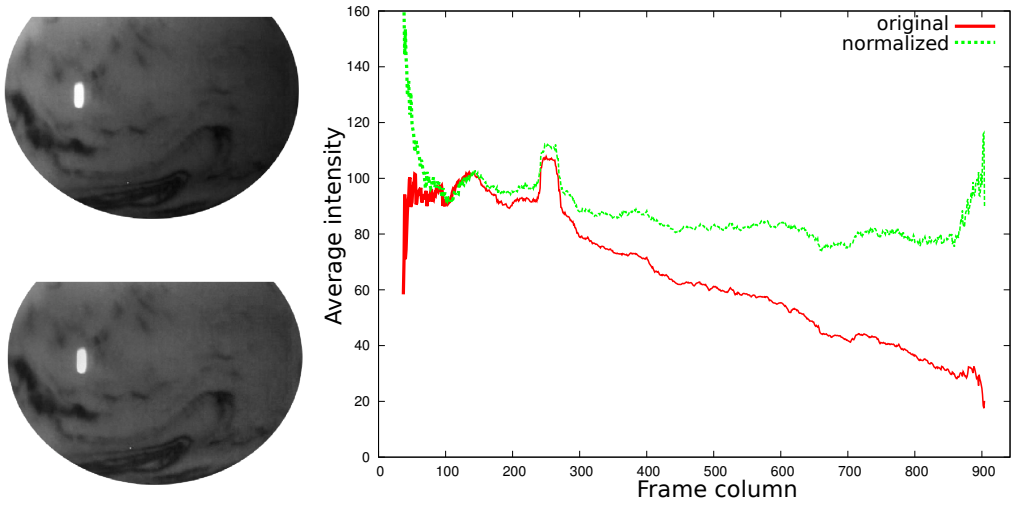


Figure 3.4: Normalization by contrast and luminosity correction. From left to right: original frames, preprocessed frames, and average intensity by frame columns for both options. The left side of the original frames has higher intensity values than the right side, whereas the average intensity is uniform in the normalized frames.

3.2 Break-Up Segmentation

The break-up area is computed from the intensities of the green component of the normalized tear film frames. Each tear film video presents variations in color and lightness related to biological characteristics and the amount of fluorescein instilled, so not all the SOIs have the same intensity levels. Furthermore, the dark pixels at the break-up vary in a range close to zero according to lighting conditions, but not exactly zero. Moreover, fluorescein evaporates and break-up areas become darker as time passes by, until reach the SRF. For these reasons, the frame intensities are analyzed in each SOI in order to de-

termine a break-up threshold, that is, the maximum intensity for a pixel to be considered as part of the break-up. On one hand, the SRF of each SOI is analyzed for getting a break-up threshold for computing the BUT measurement. The SRF provides a representation of intensities before the break-up, so it is suitable for detecting the first appearance of dark areas. On the other hand, the intensities of ERF are analyzed in order to characterize the break-up areas at the end of the SOI. Since the intensity values are lower in this frame due to the evaporation of the fluorescein, the break-up threshold computed in the SRF could not be adequate.

Two alternatives have been explored for analyzing the frame intensities. In the first approach, t_b is computed from the cumulative histogram of the frame, using the following equations:

$$F(x) = \frac{\sum_{i,j} I_x(i,j)}{n_{ROI}} \tag{3.6}$$

$$I_x(i,j) = \begin{cases} 1 & if \quad I_N(i,j) < x \\ 0 & otherwise \end{cases} \tag{3.7}$$

where n_{ROI} is the number of pixels in the adjusted ROI, x represents each gray level, and I_N is the normalized frame. Therefore, the threshold t_b corresponds to the largest value of x included in the percentage p_b of the pixels analyzed, as follows (see Fig. 3.5):

$$t_b = x \longrightarrow F(x) = p_b \tag{3.8}$$

Pixels with values below t_b are considered as black and belonging to the break-up area. Thus, the break-up areas can be segmented, as shown in Fig. 3.6.

The second approach applies a multilevel thresholding [Arora et al., 2008]

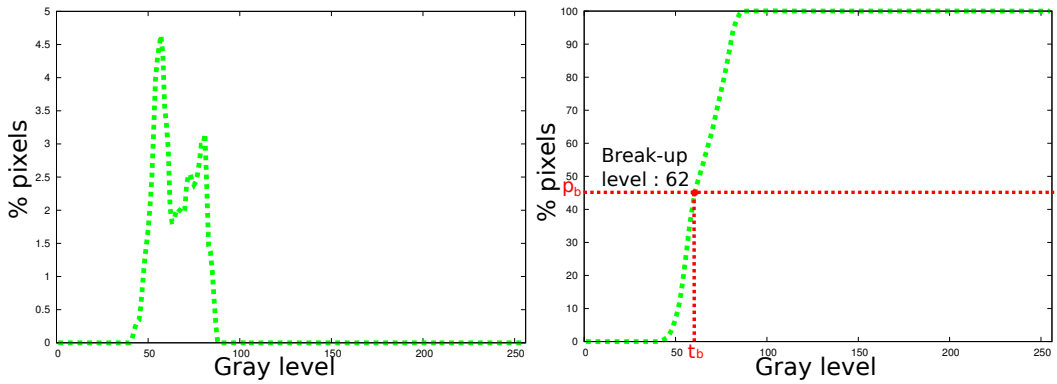


Figure 3.5: Left: histogram of the start reference frame SRF. Right: Cumulative histogram of the start reference frame SRF. The threshold t_b is obtained as the highest gray level of a percentage p_b of the darkest pixels.

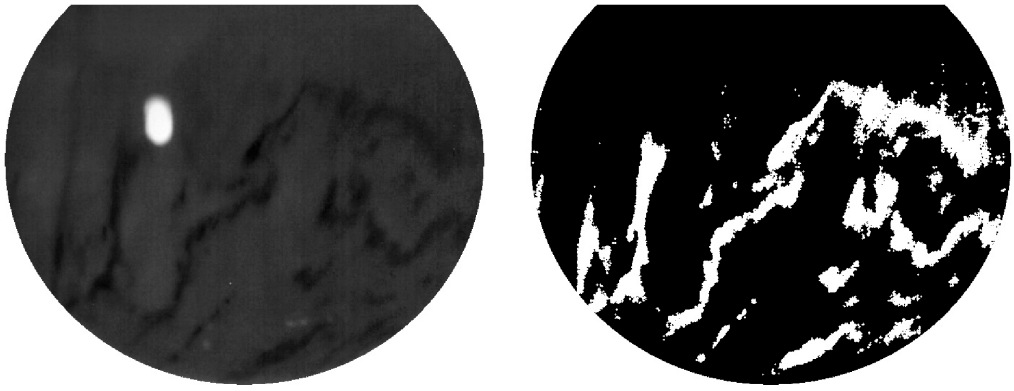


Figure 3.6: Break-up segmentation. Left: original frame at the end of a SOI. Right: break-up areas detected by thresholding using t_b.

from the statistic distribution of the frame intensities. This method uses the mean and the variance of the image to find optimum thresholds for segmenting the image into multiple levels. The algorithm 1 summarizes the steps for multilevel thresholding. It takes initially the range which include all the intensities, that is, from 0 to 255 and computes the mean μ and the standard deviation σ of all the pixels in this range. From these values, the limits t_1 and t_2 for a new sub-range are calculated as follows:

$$t_1 = \mu - k\sigma \qquad t_2 = \mu + k\sigma \tag{3.9}$$

where k is a free parameter. The algorithm is applied recursively on sub-ranges computed from the previous step so as to find a threshold and a new sub-range for the next step, until the number of chosen thresholds is reached.

Algorithm 1 Multilevel thresholding

$n \leftarrow numberofthresholds$
$end \leftarrow \frac{n}{2} - 1$
$a \leftarrow 0$
$b \leftarrow 255$
repeat
  $R \leftarrow [a, b]$
  $\nu \leftarrow mean(R)$
  $\sigma \leftarrow std(R)$
  $T1 \leftarrow \nu - k \cdot \sigma$
  $T2 \leftarrow \nu + k \cdot \sigma$
  $levels \leftarrow add([T1, T2])$
  $a \leftarrow T1 + 1$
  $b \leftarrow T2 - 1$
until $n \geq end$

Thus, pixels with intensity values in the intervals defined, are set to the respective weighted means of their values. Fig. 3.7 shows two examples of multilevel thresholding using different number of levels. Increasing the number of classes provides a better definition of the input frame, but reached certain number of levels, the addition of classes hardly varies the final segmentation.

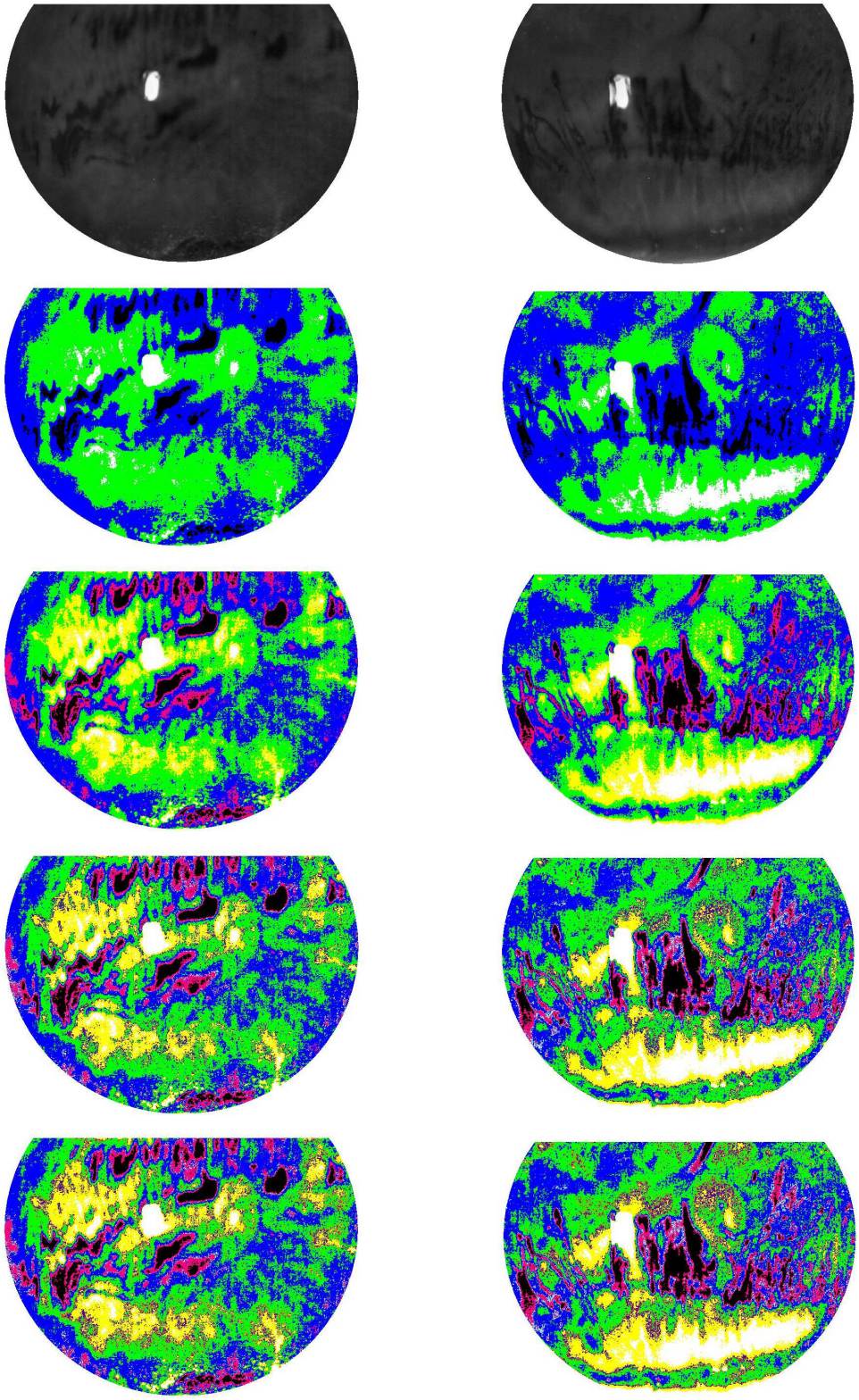


Figure 3.7: Multilevel thresholding with different levels for two SRF (one per column). From top to bottom: Original frame, 4-level, 6-level, 8-level, and 10-level thresholding.

The threshold t_b is computed from l_b, one of the lowest levels obtained in the multilevel thresholding, since they represent the darkest areas of the frame. Thus, these levels are analyzed in order to get the best threshold to segment the break-up areas. The threshold t_b corresponds to the upper limit of the selected level l_b. Figure 3.8 shows the break-up segmentation by selecting different levels l_b as threshold t_b. The lower levels segment the darkest areas whereas the upper levels segment almost all the image.

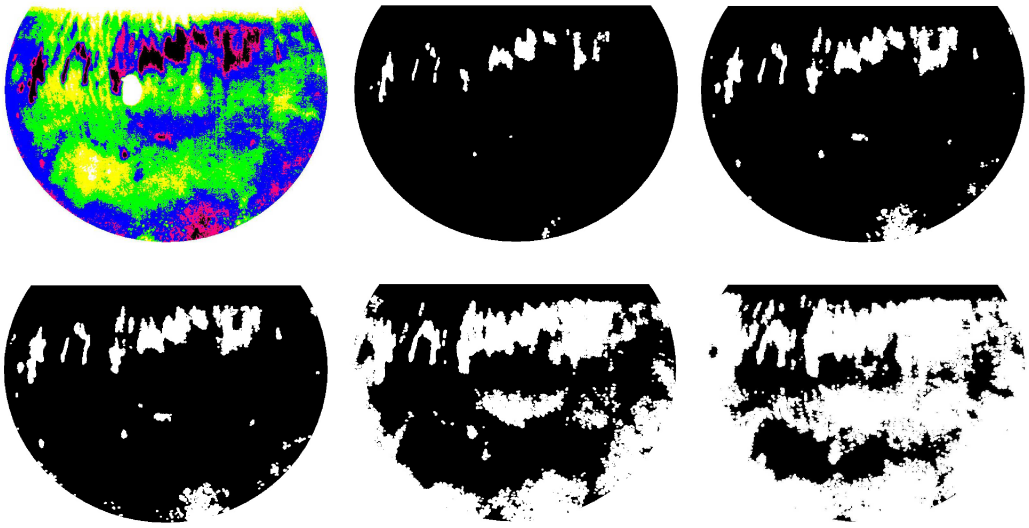


Figure 3.8: Break-up segmentation by multilevel thresholding. The top left frame is the original result of $6 - level$ thresholding and the rest are the break-up segmentation with t_b obtained with values of l_b from 1 to 5.

3.3 Break-Up Time (BUT)

The BUT is computed as the time elapsed since the last blink until the tear film break-up, that is, the appearance of dark areas related to a thinning of the tear film on the surface of the eye. In order to detect the emergence of these points, the percentage of break-up pixels is computed for each frame, and then, the evolution of these percentages is examined through the SOI.

Therefore, the percentage of break-up is obtained by applying the threshold t_b at each frame k in the SOI, as follows:

$$BU_k = \frac{\sum_{i,j} T_k(i,j)}{n_{ROI}} * 100 \tag{3.10}$$

$$T_k(i,j) = \begin{cases} 1 & if \quad I_{K_N}(i,j) < t_b \\ 0 & otherwise \end{cases} \tag{3.11}$$

where n_{ROI} is the number of pixels in the adjusted ROI, x represents each gray level, and I_{K_N} is the normalized frame k. Then, an evolution curve is computed from the differences between the break-up percentages of each frame k and the break-up percentage of the SRF frame.

Small variations produce curves with irregular slopes, so a curve fitting is performed to discard these fluctuations. Different families of curves, such as polynomial functions with different order or exponential functions, were considered to approximate the original evolution curve. A second order polynomial function has been selected for this adjustment since it provides a good fit with a low computational cost. More complex functions provide similar results of fitting but present higher computational cost.

In some cases, the gradient of the curve is zero because the tear film does not break in the interval, as shown in Fig. 3.9 (top). On the contrary, if there is a measurement, the percentage of black increases with the time since the fluorescein is not regenerated, as shown in Fig. 3.9 (bottom).

The BUT is computed from the curve that represents the evolution of the percentage of segmented break-up, from the frame where the eye is fully open up to the frame with the final blink. In order to determine the BUT measurement, the threshold t_e is obtained from the percentage p_e of the total height of the evolution curve. The BUT measurement is computed as the time elapsed from the beginning of the SOI until the curve exceeds this threshold

t_e, as shown in Fig. 3.9 (bottom).

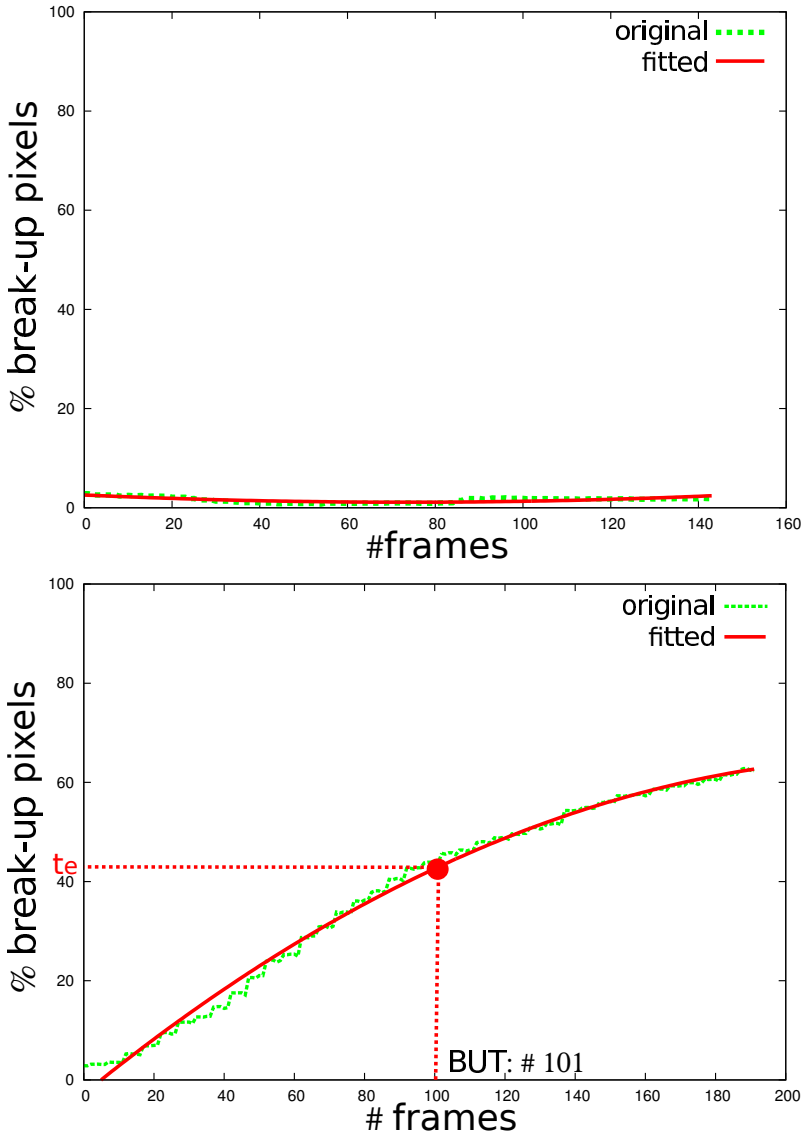


Figure 3.9: Evolution curves fitted by a second order polynomial function. Top: flat curve due to a stable tear film. Bottom: the black area increases with time and the BUT is detected when the curve exceeds the threshold t_e.

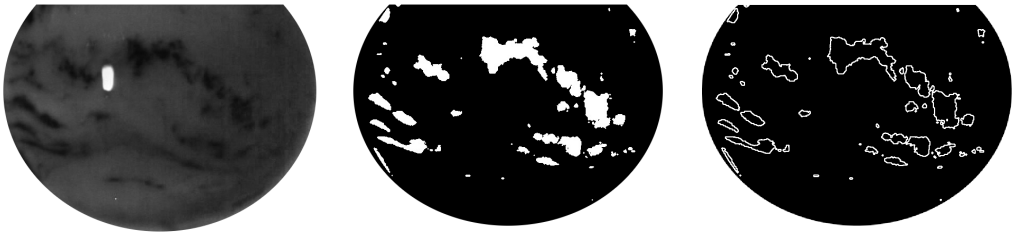
3.4 Break-Up Classification

The first break-up appears as a dark area which could increase its size with time presenting different rupture patterns. Thus, break-up could expand in one direction, forming a line, or in several directions in a homogeneously or heterogeneously way. Three different rupture patterns defined as streaks, dots, and pools have been identified in the literature [Bitton and Lovasik, 1998]. Figure 3.10 shows different examples of each pattern extracted from different SOIs. The streak rupture pattern presents a linear shape while the dot rupture pattern has a circular morphology. The pool rupture pattern is characterized as a disturbance of the tear film which conforms an irregular region that has neither a linear nor a circular shape.

Figure 3.10: Streak, dot and pool rupture patterns, in this order.

In this step, the shape of the segmented break-up areas is analyzed in order to classify them into the different rupture patterns. To this end, the ERF is selected in each SOI and the break-up areas are extracted by thresholding. In order to discard noise, the rupture zones are smoothed by applying morphological operations of opening and closing [Dougherty, 1992]. For analyzing the shape of the break-up areas, the break-up contours are computed (see Fig. 3.11, c) [Suzuki and Abe, 1985]. The algorithm for extracting the contours takes as input the thresholded image in which the break-up areas are repre-

sented by pixels with value 1 and the rest of the image by pixels with value 0. Starting with the upper left corner, the algorithm scans the image until it finds the first pixel with value 1, defined as starting point. Candidates to contour points are those pixels with value 1 and which have some pixel with value 0 in its neighborhood. Thus, the connected pixels which satisfies the condition of contour points are followed until the starting point is reached. This set of pixels with value 1 defines the contour of a break-up area. After this, the scan resumes, looking for new starting points. When the scan reaches the lower right corner of the frame, the algorithm stops.



(a) (b) (c)

Figure 3.11: (a) Last frame of the SOI. (b) Break-up areas. (c) Break-up contours used for break-up classification.

Break-ups with the same rupture pattern show a high variability so there is no general models to characterize them. For this reason, the break-up zones are classified by analyzing their morphological features. For this purpose, spatial and central moments are computed for each contour [Hu, 1962]. Moments provide a geometrical meaning by different parameters which allows to extract different features for each contour [Reed Teague, 1980; Nunes et al., 2010]. The spatial moments are calculated from the following equation:

$$m_{p,q} = \sum_{i,j} i^p j^q \, T_k(i,j) \tag{3.12}$$

where $T_k(i,j)$ corresponds to the thresholded image, where the background

pixels are 0 and the pixels belonging to the contour are 1. The moments are usually classified by the order, which depends on the indices p and q. Thus, the sum $p + q$ is the order of the moment $m_{p,q}$. The zero order moment $m_{0,0}$ describes the area A delimited for the contour. The first order moments $m_{1,0}$, and $m_{0,1}$, contain information about the center of gravity of the contours $(\bar{i}, \bar{j})$, from which can be derived the central moments :

$$mu_{p,q} = \sum_{i,j} (i - \bar{i})^p (j - \bar{j})^q I(i,j) \tag{3.13}$$

This way, a shape descriptor is built with the features extracted from the moments of each contour. On one hand, according to the shape of the different rupture patterns, the features should distinguish between linear and circular morphologies. This way, dot and streak patterns can be discriminated. For this purpose, the following features are computed:

- **Axis ratio r_{ab}.** The main inertial axes of the object correspond to the semi-major and semi-minor axes a and b of the ellipse which can be used as a approximation of the considered contour. The main inertial axis are those axis, around which the contour can be rotated with minimal (major semi-axis a) or maximal (minor semi-axis b) inertia. The main inertial axis a and b, shown in Fig. 3.12, can be derived from the second central moments:

$$a, b = \sqrt{\frac{1}{2}(mu_{2,0} + mu_{0,2} \pm \sqrt{(mu_{2,0} - mu_{0,2})^2 + 4mu_{1,1}^2})} \tag{3.14}$$

  r_{ab} is the ratio between the semi-major and the semi-minor axis. This ratio gives an idea whether the contour is more or less elongated.

$$r_{ab} = \frac{a}{b} \tag{3.15}$$

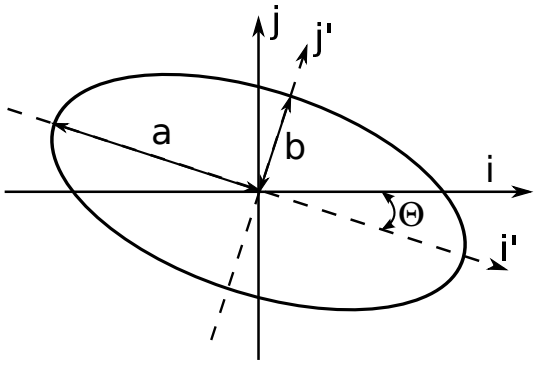


Figure 3.12: Each contour can be approximate by an ellipse. The semi-axis and the orientation provides useful features to characterize the shape of the contour.

- **Roundness κ.** The roundness κ is computed from the perimeter ρ and the area A:

$$\kappa = \frac{\rho^2}{A} \tag{3.16}$$

  If the contour is a circle κ is equal 1, for other objects is greater than 1.

- **Eccentricity ϵ.** The eccentricity ϵ can directly derived from the semi-major and semi-minor axes a and b of the contour:

$$\epsilon = \frac{\sqrt{a^2 - b^2}}{a} \tag{3.17}$$

  The eccentricity ϵ can have values from 0 to 1. Values of 0 are related to a perfectly rounded contour and values of 1 correspond to a line shaped contour.

However, pool patterns are regions with variable morphology, which can be small or large, with irregular boundaries, and can present an elongated or circular global distribution. Thus, on one hand, the area is used as a feature to distinguish between dots and small circular pools.

- **Area** A. The number of pixels of the object bounded by the contour, which is defined by the spatial moment of zero order:

$$A = m_{0,0} \tag{3.18}$$

On the other hand, the convex hull area, shown in Fig. 3.13, is related to the number or size of concavities in the contour and can be combined with the perimeter or area to get an indicator of the contour roughness. Thus, the next features are added to the shape descriptor:

Figure 3.13: Convex hull area. Red line represents the smallest convex polygon that can contain the break-up contour.

- **Convexity** C It is defined as the ratio of perimeters of the convex hull ρ_h over the perimeter of the original contour ρ.

$$C = \frac{\rho_h}{\rho} \tag{3.19}$$

- **Solidity** S It describes the extent to which the shape is convex or concave, and is defined by the relation between the area A of the break-up region and the corresponding convex hull area H. The solidity of a convex shape is always 1.

$$S = \frac{A}{H} \tag{3.20}$$

All these metrics form up a descriptor that is used as the input of a classifier to decide the final type of each break-up area. Figure 3.14 shows an example of rupture patterns identified in a SOI.

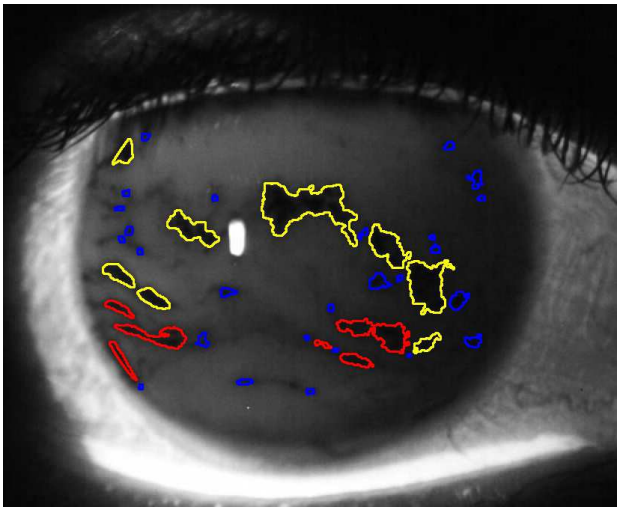


Figure 3.14: Break-up classification. Dots, streaks, and pools are depicted in blue, red, and yellow, respectively.

3.5 Break-Up Evolution

Another interesting break-up feature, besides the rupture patterns, is the evolution from the break-up frame until the end of the SOI. For this purpose, n_d frames located every 0.5 seconds (for example, 15 frames for a 30 fps video) after the first break-up until the last blink are selected. For each of these frames, the area BU_k of the tear film break-up is obtained from the percentage of break-up pixels in the ROI, as follows:

$$BU_k = \frac{n_{bu}^k}{n_{ROI}} * 100 \tag{3.21}$$

where n_{bu}^k is the break-up pixels of the frame I_k and n_{ROI} is the total number of pixels of the ROI, which is the same for all the frames in the SOI. The growing

rate r_b is computed from the first break-up until the SRF by calculating the mean of the differences between the break-up areas segmented in the selected frames, as follows:

$$r_b = \frac{\sum_{k=1}^{n_d} \Delta BU_k}{n_d} \tag{3.22}$$

$$\Delta BU_k = BU_{k+fr} - BU_k \tag{3.23}$$

where k is the current frame, and n_d is the number of frame selected every fr frames from $k = 1$ until the end of the SOI. This analysis provides additional information to the BUT test since the break-up size and its evolution are quantified from the first appearance until the later blink. Therefore, this assessment allows to characterize if the break-up is small or large and if it evolves fast or slowly. Figure 3.15 shows two examples where the break-up dynamics graphs distinguish between a large break-up with slow and short increase, and a small break-up with massive expansion.

3.6 Local Analysis

The break-up location is related to specific aspects of the tear film that could affect to dry eye severity [King-Smith et al., 2009]. On one hand, the central cornea is much more sensitive than the periphery, so tear film break-up at this zone should affect vision as a result of tear film changes over the pupil [Liu et al., 2006]. On the other hand, different SOIs with break-up at the same location could indicate that this area is weakened. Thus, break-up location is another relevant break-up feature, so a local analysis would provide additional information.

In order to extend the global methodology to the local analysis, the ROI is divided into five equally sized zones according to the Cornea and Contact Lens Research Unit (CCLRU) standards [Terry et al., 1993a,b], and the break-up

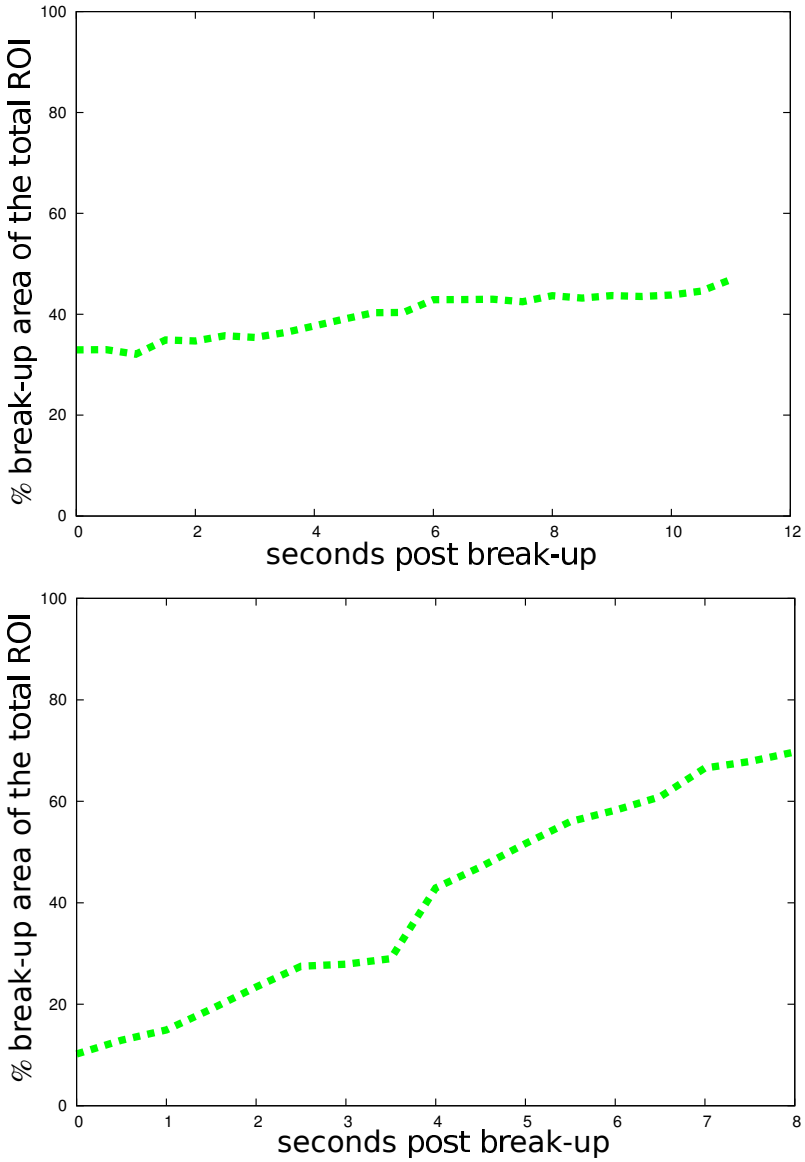


Figure 3.15: Break-up dynamics in two different SOIs. Top: the first break-up represents the 33% of the whole ROI and evolves at a growing rate of 1.75%/s until reach the 47% of the whole ROI. Bottom: the first break-up covers the 11% and the growing rate is of 7.5%/s, occupying almost the 70% of the ROI at the end of the SOI.

assessment is performed separately for each one. Figure 3.16 shows the grid applied over the ROI to extract the local areas and perform the break-up analysis on them in the same manner as the global analysis is performed on the whole ROI.

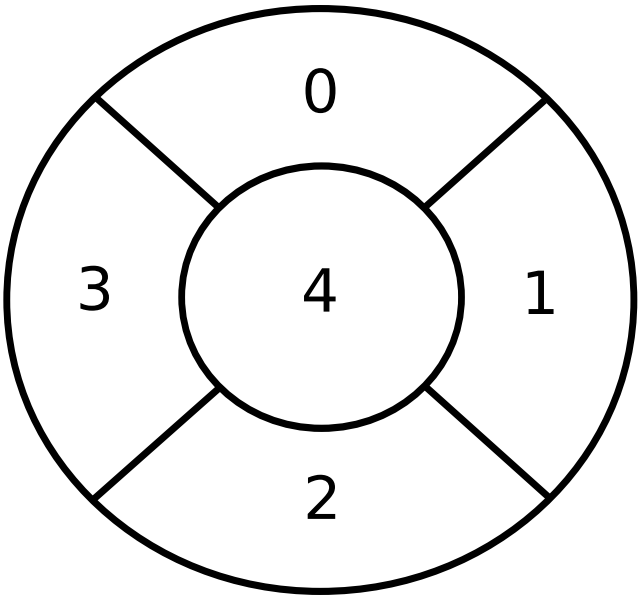


Figure 3.16: Local zones from CCLRU standards.

Therefore, for each local zone a specific break-up threshold is computed for segmenting the break-up areas and building the evolution curve for that zone. This way, five evolution curves are built, one for each local zone, as shown in Fig. 3.17. Another specific thresholds to get the BUT measurement are computed from a percentage of the maximum height of each local curve. Thus, a local BUT measurement is obtained for each local zone as the time elapsed until the curve exceeds the corresponding local threshold. Finally, the global BUT measurement is obtained from the earliest local BUT, since it represents the first appearance of dark areas, that is, the tear film break-up. Moreover, the local evolution curves have more quality since they are representing specific areas focused on the break-up zone. The maximum of

each evolution curve gives an idea about the amount of break-up region in this zone. Similarly, the growth of the curve is related to the break-up dynamics. Furthermore, each of the break-up features evaluated in the global analysis can be performed separately in each local zone, obtaining a more specific break-up characterization.

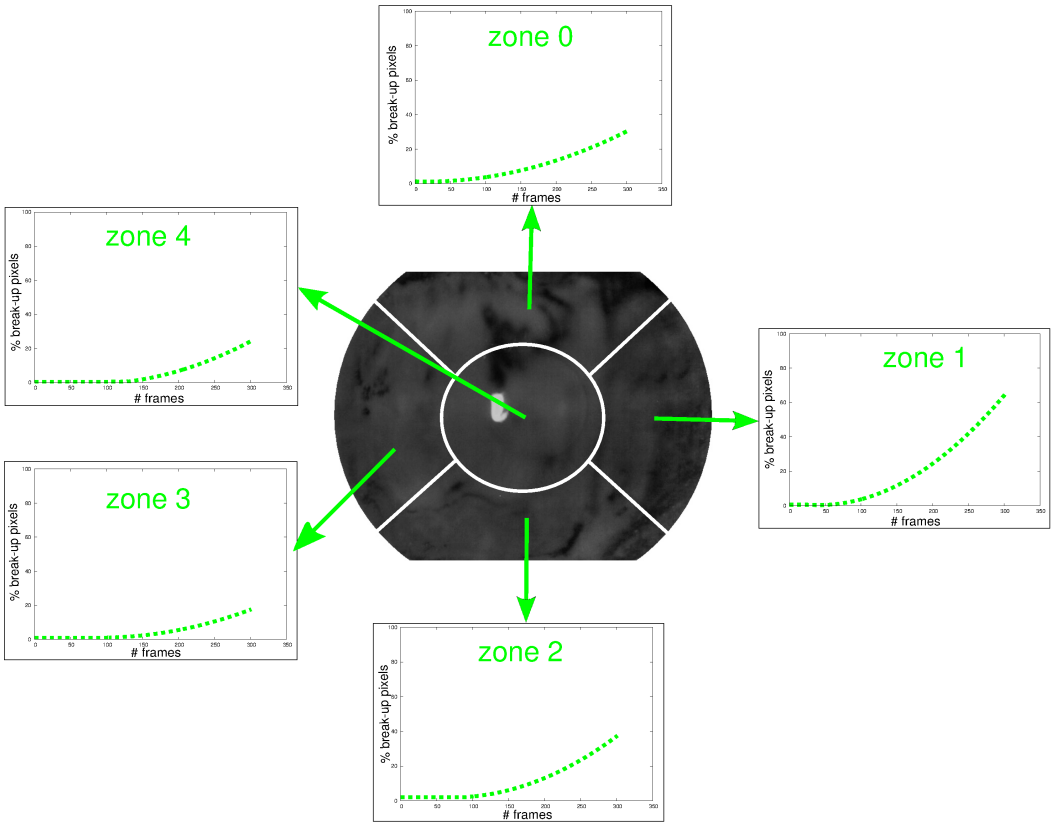


Figure 3.17: Fitted evolution curves for the five local zones. The height of the curve represents the amount of break-up at this zone, while the slope is related to break-up dynamics.

Besides the BUT measurement, the local analysis provides useful information about the zone where the break-up occurs. This way, a very localized break-up that affects to a particular zone, can be distinguished from a break-up distributed by several zones or a massive break-up which covers the whole ROI. Figure 3.18 shows the local evolution curves in three SOIs.

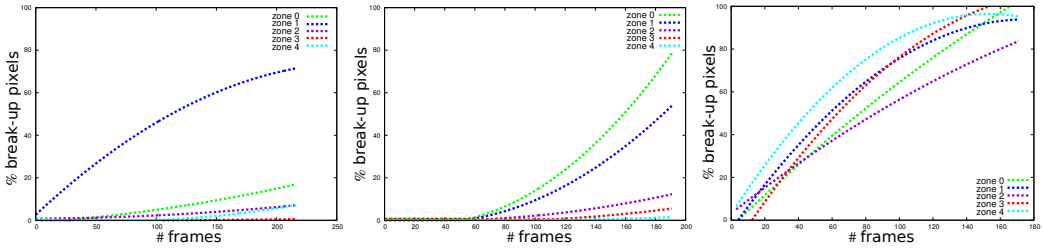


Figure 3.18: Local evolution curves in different SOIs. In the first case, the break-up occurs mainly at zone 1, while the growing of the remaining zones is low. In the second case, the break-up is mainly distributed over the zones 0 and 1, whereas, in the third case, a massive break-up occurs over the whole ROI.

If the break-up occurs in the same location in different trials, it could indicate that this area is weakened. The local analysis allow to detect such situations when the local evolution curves follow a similar pattern. Figure 3.19 shows the evolution curves for three different SOIs of the same patient. In this cases, the break-up occurs in the same area of the eye and the local evolution curves present a similar evolution.

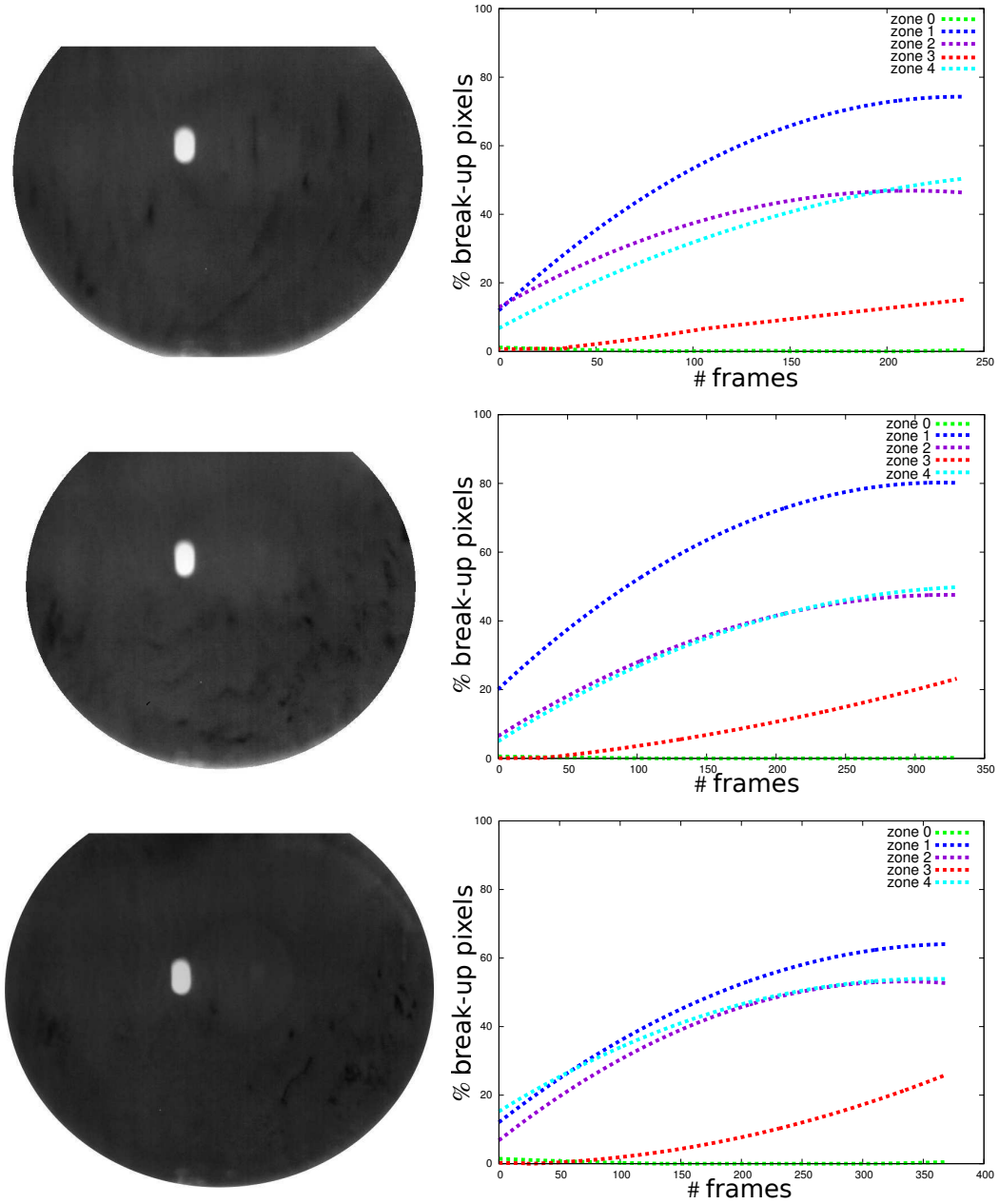


Figure 3.19: Local evolution curves for 3 SOIs of the same tear film video. The break-up is mainly distributed over the zones 1, 2, and 4, so the corresponding evolution curves present a high growing. The zone 3 has a small break-up, which is reflected in the slightly increasing curve of this zone. The zone 0 does not break-up, so the local evolution curve is flat for all the SOIs.

Chapter 4

Results

The experiments performed to test the methodology for the break-up assessment are described in this chapter, summarizing the main obtained results.

In order to validate the methodology, the steps are analyzed separately. Therefore, the SOI location, the extraction and adjustment of the ROI, the break-up segmentation, the accuracy of the BUT detection as well as the break-up classification are validated in relation to the information provided by the experts. Moreover, a time analysis is conducted to check if the methodology can approximate to real time use.

4.1 Materials

Our methodology has been tested on tear film videos provided by the Servizo de Optometría directed by Dra. Eva Yebra-Pimentel at the Facultade de Óptica e Optometría of the Universidade de Santiago de Compostela. These videos have been recorded with a Topcon DV-3 camera attached to a Topcon SL-D4 slit lamp. The dataset consists of 18 videos from healthy patients with ages ranging from 19 to 33, varying from very dry eye to no visible dryness, as shown in Fig. 4.1.

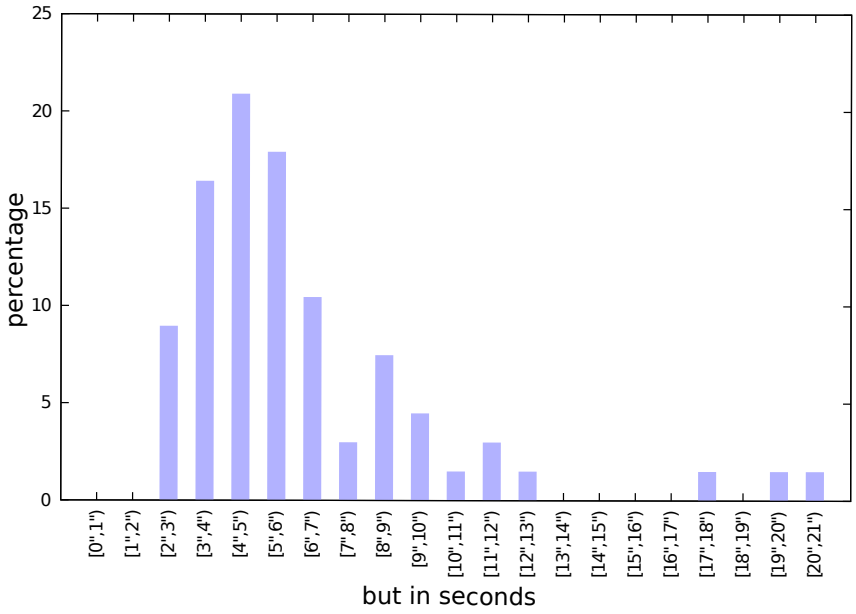


Figure 4.1: BUT measurements obtained from the mean of the annotations provided by four experts in the dataset.

These videos have been manually annotated by four different experts for validating each step of the methodology. Table 4.1 details the data used in each experiment.

Table 4.1: Data used for validating the different steps of the methodology.

Evaluation	Dataset	Dataset selection
SOI Location	18 videos	113 SOIs (63 with BUT/50 without BUT)
ROI Extraction	63 SOIs	150 random frames
Break-up Segmentation	63 SOIs	140 random frames
BUT Measurement	18 videos	63 SOIs with BUT
Break-up Classification	63 SOIs	ERF frames (33 pools, 38 streaks, 25 dots)

The proposed methodology has been implemented in C++ using the Open Computer Vision library (OpenCV) for performing some video and image processing operations. The development and evaluation have been conducted in a GNU/Linux operating system running on an Intel Pentium Quad processor at 2.33GHz and 4 GB of RAM.

4.2 Analysis of SOI Location

The SOI location consists in identifying the sequences of frames between consecutive blinks where the BUT test can be performed. This step has been validated by comparing the SOIs detected by the system with the SOIs delimited by the experts in the full dataset of 18 tear film videos. The different SOIs are delimited using the differences of frame intensities according to Eq. 2.5. Table 4.2 shows the performance of SOI location obtained with values of f from 1 to 3.

Table 4.2: Performance of SOI detection with different values of the parameter f.

	% located SOIs	Invalid sequences
f = 1	92.32	7 (semi-blinks, lamp off)
f = 2	97.14	1 (semi-blink)
f = 3	97.14	1 (semi-blink)

As can be seen, using $f = 1$ some SOIs are missed as well as some invalid sequences are detected, whereas including differences between non consecutive blinks provides better results, since this sum emphasizes the intensity changes produced during a blink. This way, with $f = 2$ nearly all the SOIs are located and almost no invalid sequence is included. However, with $f \geq 2$ the improvement is negligible, so f is set to 2 since the computational complexity is lower and this value works well detecting SOIs with BUT measurement as well as discarding invalid sequences which contains semi-blinks or periods where the lamp is off.

4.3 Evaluation of ROI Extraction

The ROI extraction consists of a stage for identifying the position and size of the iris, and then, another stage for discarding outside elements such as eyelids and eyelashes. In order to validate the ROI extraction step, each stage is analyzed separately. A random selection of 150 frames has been considered for both evaluations. Each frame of this dataset has been manually annotated by an expert. The iris was marked by an ellipse within each frame, and then, the top and the bottom of this ellipse were cropped discarding the outer elements such as eyelids or eyelashes. The manual results provided by the expert were compared to the automatic results extracted by the system.

First, the accuracy of the automatic iris segmentation has been tested using two statistical measures, the sensitivity and the specificity, which compute the proportion of iris pixels correctly detected as well as the non iris pixels correctly discarded. An ideal segmentation technique should detect all the iris pixels and should avoid the background areas. Therefore, the True Positives (TP) are defined as the iris pixels correctly detected, the True Negatives (TN) are the background pixels correctly discarded, the False Positives (FP) are the background pixels detected as foreground areas whereas the False Negatives (FN) are the iris pixels detected as background areas, as shown in 4.2.

The sensitivity and the specificity are computed from the previous values as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.1}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4.2}$$

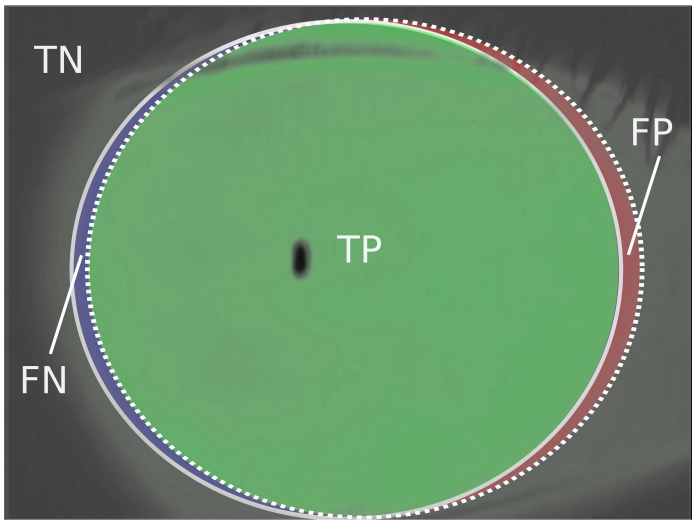


Figure 4.2: Validation of ROI segmentation in terms of sensitivity and specificity. The dotted lines represent the ROIs detected by the system and the continuous lines are the manually delimited ROIs.

This way, the sensitivity measures the proportion of iris pixels correctly detected, whereas the specificity measures the proportion of background pixels discarded. If both measures have values near 100, the automatic method is able to segment the iris as the expert does. Table 4.3 shows the results obtained from the mean of these measures. These results represent that the most part of the iris is segmented and slight variations in relation to the experts are due to the blurred iris boundaries.

Table 4.3: Performance for the iris identification in terms of sensitivity and specificity.

	Iris identification
Sensitivity (%)	95
Specificity (%)	90

Once the position and size of the iris are identified, an adjustment is performed in order to discard outside elements such as eyelids and eyelashes which could disrupt the results. This adjustment consists in reducing the radii, and cropping the top and the bottom of the ellipse. The parameters related to this

adjustment are set by building ROC curves and selecting the best values on the basis of sensitivity and specificity criteria [Fawcett, 2006; Powers, 2007]. First, the parameter γ (see Eq. 2.16) which represents the percentage of the original radius length is set by comparing the manual ellipses marked by the expert to the automatic ellipses extracted by the system. This parameter reduces the ellipse radii in order to discard the noise related to the iris boundaries. To this end, values from 80% to 95% are considered for building the ROC curve. Sensitivity and specificity are obtained from Equations 4.1, 4.2, where the TP, TN, FP and FN for each case are computed in the same way as in the previous validation stage. The parameters α and β (see Eqs. 2.14, 2.15) are related to the top and bottom cropping, representing the percentage of the maximum number of edge points found in the upper and lower half of the image, respectively. Thus, this validation is performed considering each half of the frame separately, with α and β taking values from 5% to 95%, as shown in Fig. 4.3.

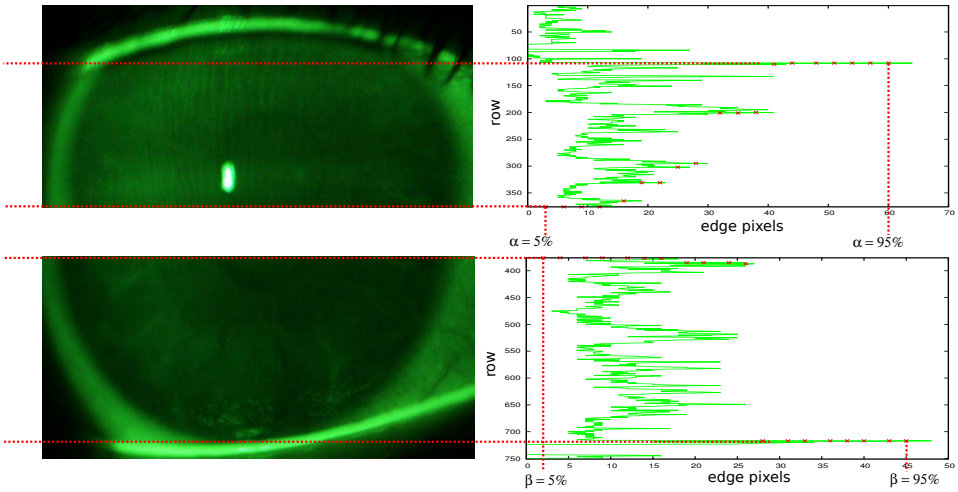


Figure 4.3: Top and bottom cropping with α and β taking values from 5% to 95% of the maximum number of edge points found in the upper and lower half of the image.

The manual adjustment performed by the expert is correlated to the automatic adjustment provided by the system, and TP, TN, FP and FN are computed in each half of the frame, as shown in Fig. 4.4.

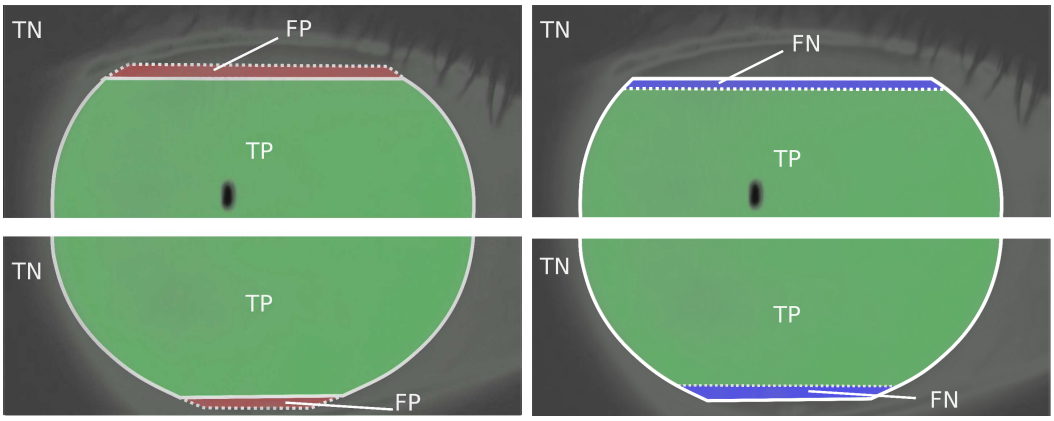


Figure 4.4: Validation of ROI adjustment in terms of sensitivity and specificity. The top and bottom cropping are separately evaluated. The dotted lines represent the automatic cropping performed by the system and the continuous lines are the cropping selected by the expert.

Figure 4.5 shows the ROC curves for the parameters α, β, and γ. The best parameters would yield a point in the upper left corner of the ROC space, representing 100% sensitivity and 100% specificity. In order to get a compromise with these two measures, a line from the upper left point to the opposite corner is drawn, and the parameters are set to the values of the ROC curve crossed by this line. Thus, α is set to 70%, β is set to 90%, and γ is set to 95%. The results of sensitivity and specificity for these parameters are shown in Table 4.4. According to these results, most part of the ROI is segmented and the adjusted ROI is similar to the ROI manually delimited by the experts so it is appropriate for break-up assessment.

Table 4.4: Performance for the ROI adjustment in terms of sensitivity and specificity.

	Radii reduction ($\gamma = 95\%$)	Top cropping ($\alpha = 70\%$)	Bottom cropping ($\beta = 90\%$)
Sensitivity (%)	94.61	90.47	97.22
Specificity (%)	90.32	88.39	93.28

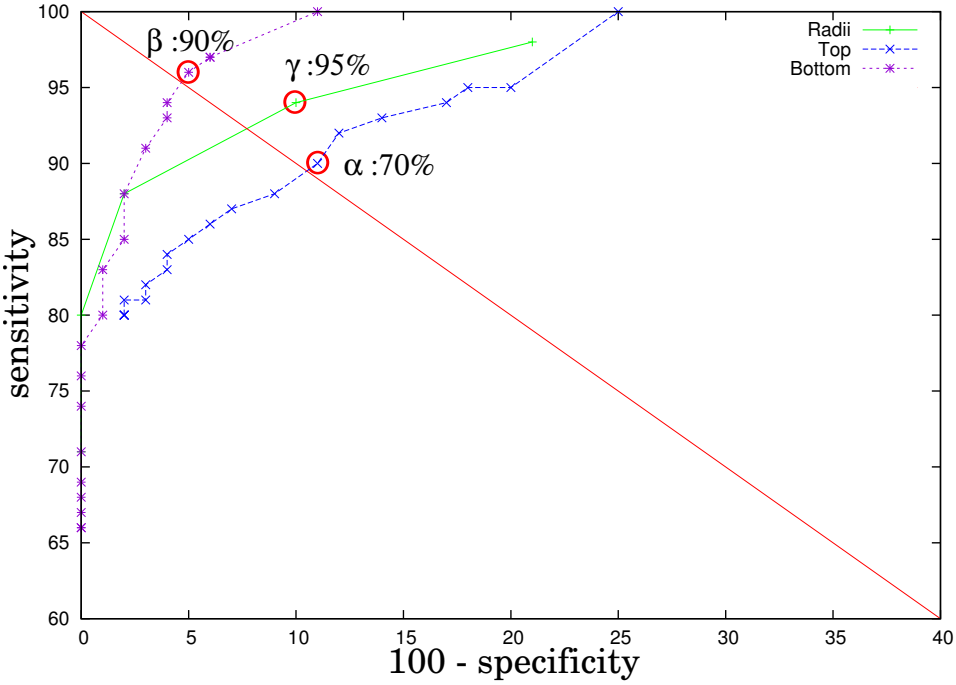


Figure 4.5: ROC curves for setting the parameters of ROI adjustment. Sensitivity and specificity were computed by comparing the ROIs extracted by the system to the manual ROIs segmented by the expert.

4.4 Evaluation of Break-up Segmentation

The break-up segmentation is computed from a break-up threshold in each SOI which represents the maximum intensity for a pixel to be considered as part of the rupture. The parameters related to the break-up threshold computation are adjusted by building ROC curves and selecting the best values on the basis of sensitivity and specificity criteria [Fawcett, 2006; Powers, 2007].

Due to the limited size of the dataset, 10-fold cross-validation is used in the experiments to assess the generalization capability [Rodriguez et al., 2010]. Therefore, the dataset composed by 140 frames manually segmented by 2 experts is divided into 10 parts and an iterative process is carried out 10 times. At each iteration, one of the parts is used as a test set, and the remaining parts are used as a training set. Finally, ROC curves are built for each 10 k-fold training set and they are analyzed to adjust each parameter. Once the best parameter values are selected, the accuracy is computed from the k-fold test sets. Considering that each pixel can be classified as a break-up pixel or a background pixel, values for sensitivity and specificity are obtained by comparing the matches pixel by pixel between the manual break-up regions marked by the experts and the automatic results provided by the system. Note that labeling the break-up areas by hand is a tedious and subjective task so that there is a remarkable disagreement between the experts, as shown in Fig. 4.6. The coincidence of the delimited break-up areas is around 60%, so the intersection between both experts is used as gold standard for the validation.

Two alternatives have been proposed for computing the break-up threshold. In the first approach, the break-up threshold t_b is obtained from the percentage p_b of the darkest pixels in the cumulative histogram of the SRF. A ROC curve is built for each k-fold training set with p_b taking values from 5% to 95%, as shown in Fig. 4.7.

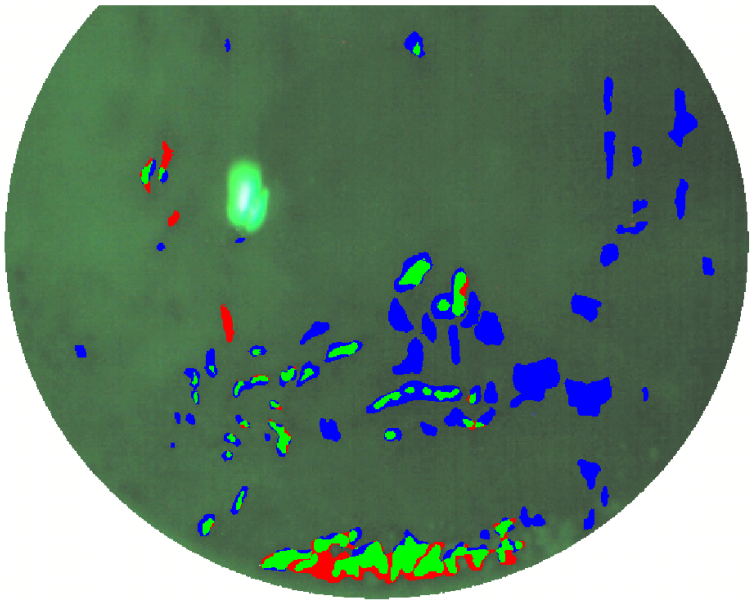


Figure 4.6: Break-up areas manually annotated by two experts in the same frame. Red regions are related to the areas marked only by Expert 1 whereas blue regions correspond to annotations made only by Expert 2. The intersection between the regions segmented by the both experts is represented in green.

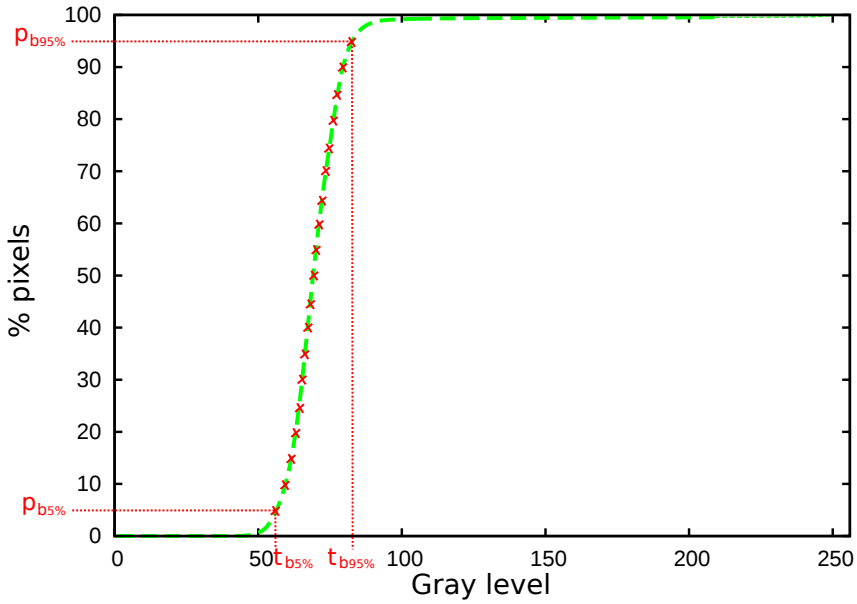


Figure 4.7: Different thresholds t_b are obtained by scanning the cumulative histogram of the SRF with percentages of p_b taking values from 5% to 95%.

In order to get a compromise between sensitivity and specificity, p_b is set to the value of the ROC curve crossed by a line from the upper left point to the opposite corner. The best values for p_b are always between 40% and 45% in the different ROC curves corresponding to the 10-fold training sets, so a mean ROC curve is computed as a representation of the training step. This ROC curve is built from the mean values of specificity and sensitivity in the 10 k-fold training sets, and is used to set the value of p_b, as shown in Fig. 4.8. Thus, p_b is set to 45% of the darkest pixels of the cumulative histogram. This value was checked in the 10 k-fold test sets providing the results showed in the Table 4.5. Appendix C.1 includes the ROC curves for each k-fold training set as well as the accuracy results of each k-fold test set.

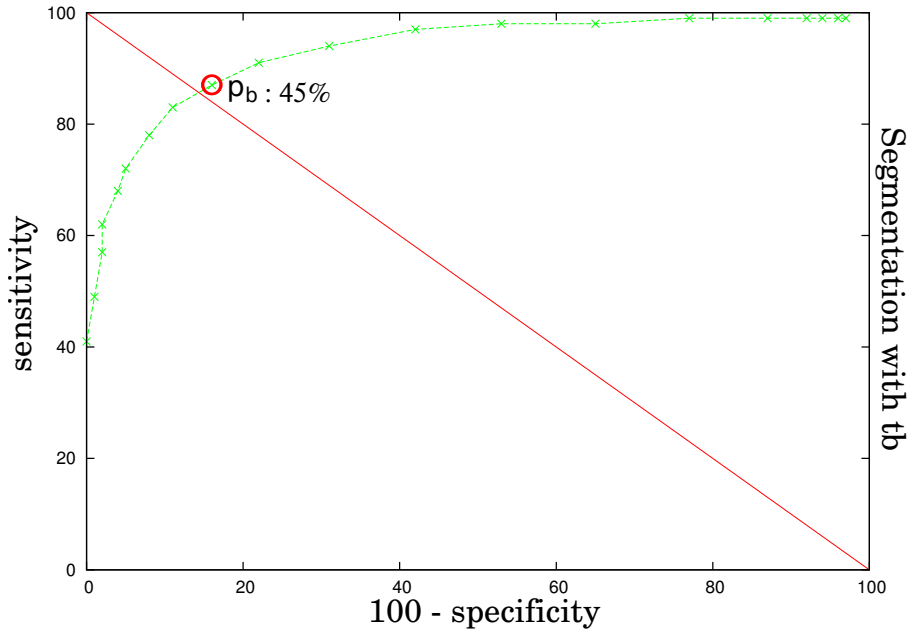


Figure 4.8: Mean ROC curve for the 10 k-fold training sets obtained with percentages from 5% to 95% of the darkest pixels. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

The second approach uses the ERF which delimits the end of the SOI. The threshold t_b is computed from upper limit of l_b, this is, one of the levels

obtained by the multilevel thresholding. In order to select the value of l_b which provides the best segmentation, all the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding have been analyzed for the different k-fold training sets. The different 10 k-fold training sets provide the same results for the best l_b in each thresholding. Thus, mean ROC curves computed from the mean values of sensitivity and specificity in the 10 k-fold training sets are used as a representation of the training iteration, as shown in Fig. 4.9. The $6-level$ presents slightly better results than the $4-level$, $8-level$, and $10-level$ thresholding. Therefore, l_b is set to level 3 in the $6-level$ thresholding, since it provides the best compromise between sensitivity and specificity. This value was checked in the 10 k-fold test sets providing the results showed in the Table 4.5. Appendix C.1 includes the ROC curves for each k-fold training set as well as the accuracy results of each k-fold test set.

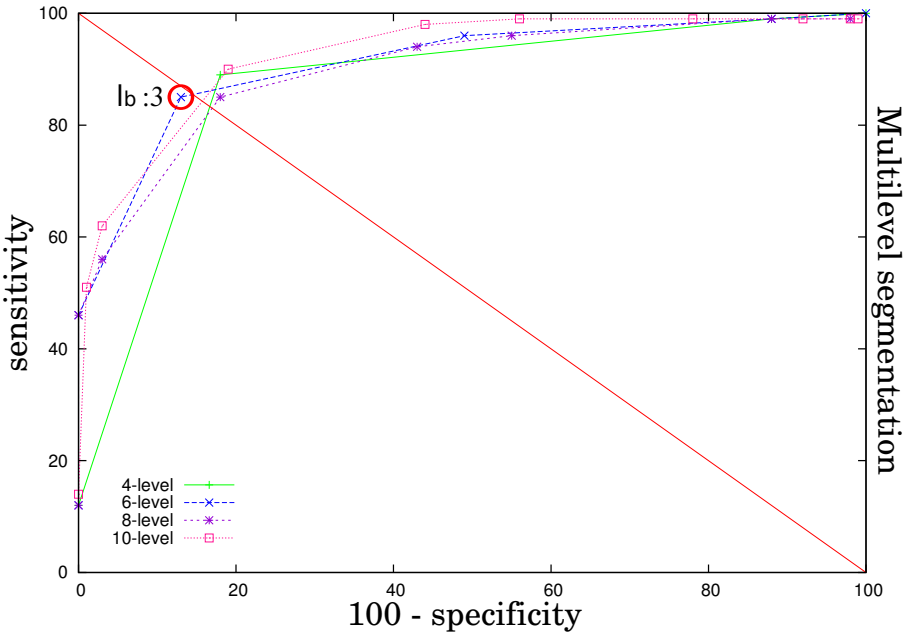


Figure 4.9: Mean ROC curve for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

Table 4.5: Accuracy (%) of break-up segmentation for the best case, the worst case, and the mean of the 10 k-fold test sets with the best values for p_b and l_b.

	Cumulative histogram ($p_b = 45\%$)	Multilevel thresholding ($6 - level, l_b = 3$)
Best case	90.94	92.84
Worst case	77.33	86.87
Mean	84.92	89.09

These approaches get acceptable results considering the high variability between the experts. The sensitivity and specificity provided by the multilevel thresholding is slightly better, so t_b is computed from the upper limit for the level l_b 3 in the $6-level$ thresholding. This alternative is used for the validation of following steps of the methodology.

4.5 Assessment of BUT measurement

The accuracy of the BUT measurement is validated in relation with the values provided by four different experts for the 18 tear film videos of the dataset. The BUT is calculated from the break-up evolution curve which represents the percentage of break-up pixels throughout the SOI. If the gradient of this curve is zero, the tear film does not break in that SOI, whereas if there is BUT measurement, the percentage increases with the time since the fluorescein is not regenerated. Therefore, before evaluating the performance of BUT measurement, the gradient of the evolution curves is analyzed to determine if there is tear film break-up at each SOI. This classification is compared to the annotations provided by the experts, as shown in Table 4.6. As can be seen, the automatic system performs well detecting the SOIs where there is BUT measurement as well as discarding the SOIs where there tear film remains stable until the ERF.

After this, the SOIs where tear film break-up is expected are used for assess-

Table 4.6: Performance (%) of classification in SOIs with BUT measurement and SOIs where tear film does not break-up.

		System	
		Measure	No measure
Expert	Measure	96	4
	No measure	8	92

ing the BUT measurement. The BUT is calculated by applying a threshold t_e to the evolution curve of each SOI. This threshold is obtained from the percentage p_e of the total height of that evolution curve. In the same way that other parameters, p_e is set by comparing the break-up frames automatically classified by the system respect to the average of the values manually annotated by the experts. The validation is performed considering that each frame can be classified as a frame with or without break-up. The frames located before the BUT are considered as frames without break-up. Similarly, the subsequent frames to the BUT are considered as frames with break-up. Therefore, TP and TN are computed from the coincidences between manual and automatic measurements of the frames classified as break-up frames and those classified as frames without break-up, respectively. On the other hand, FP and FN are obtained from the intervals between both measurements. This way, if the automatic measurement is ahead of the expert annotation, the frames in this interval are FP, as shown in Fig. 4.10 (top). However, if the automatic system detects the BUT after the expert, the frames in this interval are FN, as shown in Fig. 4.10 (bottom)

Thus, the values of sensitivity and specificity are calculated by using Eqs. 4.1 and 4.2. The ROC curve is built for each k-fold training set with p_e taking values from 5% to 95% of the total height of the evolution curve, as shown in Fig. 4.11.

The best values for p_e are always between 60% and 65% in the different k-fold iterations, so the ROC curve computed from the mean values of speci-

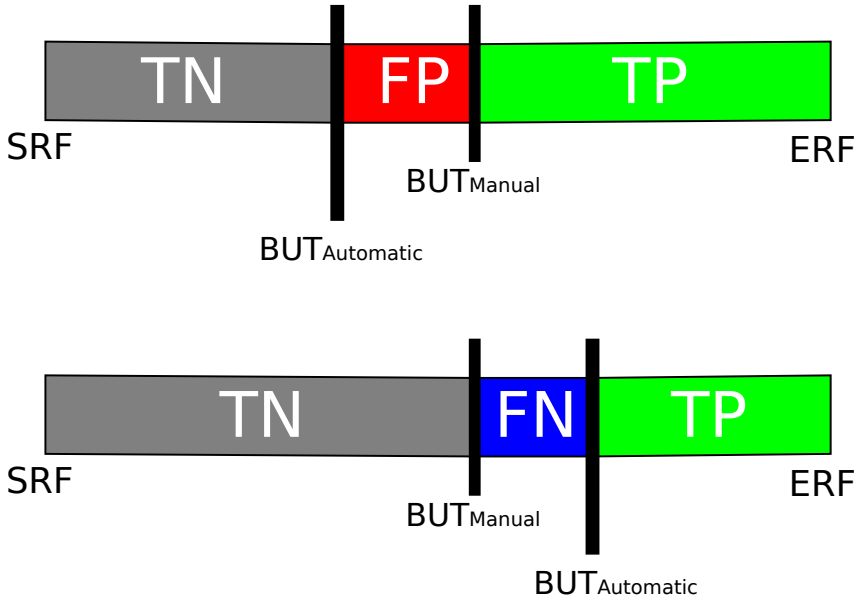


Figure 4.10: Validation of the BUT measurement in terms of sensitivity and specificity. Top: the system detects the break-up before the experts. Bottom: the system detects the break-up after the experts.

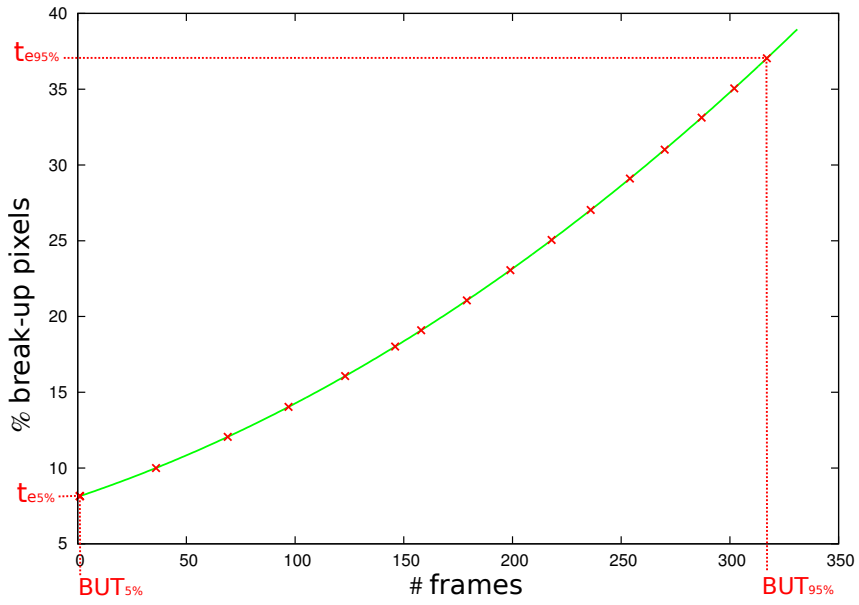


Figure 4.11: Different thresholds t_e are obtained by scanning the evolution curve with percentages p_e taking values from 5% to 95% of its maximum height.

ficity and sensitivity in the 10 k-fold training sets is used to set the value of p_e, as shown in Fig. 4.12. Thus, p_e is set to 65%, since it provides the best compromise between sensitivity and specificity. This value was checked in the 10 k-fold test sets providing the results showed in the Table 4.7. Appendix C.2 includes the ROC curves for each k-fold training set as well as the accuracy results of each k-fold test set.

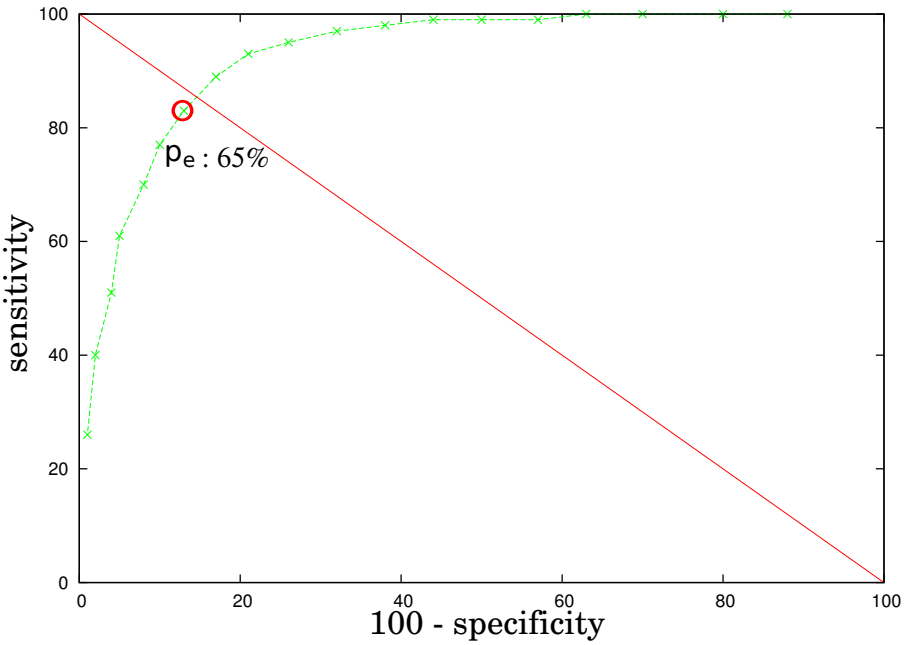


Figure 4.12: Mean ROC curve for the 10 k-fold training sets obtained with percentages from 5% to 95% of the total height of the evolution curve. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

Table 4.7: Accuracy (%) of BUT measurement for the best case, the worst case, and the mean of the 10 k-fold test sets with the best value for p_e.

	Global ($p_e = 65\%$)
Best case	90.43
Worst case	86.22
Mean	86.73

Once the parameters are set to the best values, the BUT detection is validated in terms of differences in seconds between the automatic measurements and the values provided by the experts. Prior to this analysis, the differences between the BUT measurements annotated by each expert and the average of all them have been analyzed with the aim of quantifying the variability among the experts. Figure 4.13 shows the dispersion between the expert average and each individual expert measure. The 96% of the differences among the experts are in an interval of ±2.5 seconds due to the subjectivity of this measurement.

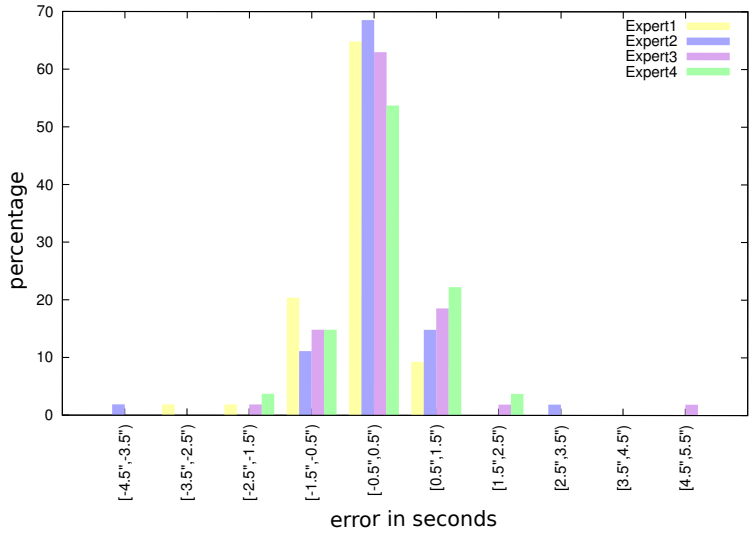


Figure 4.13: Dispersion between the expert average and the value annotated by each expert. Most of the measures have an error in an interval of ±2.5 seconds.

After that, the automatic BUT measurements have been compared to the experts' average, as shown in Fig. 4.14. The percentage of values within the range of ±2.5 seconds has been 88% and the mean deviation was 1.21 seconds. This value is in agreement with the manual detection, since it is within the same range as among the experts themselves.

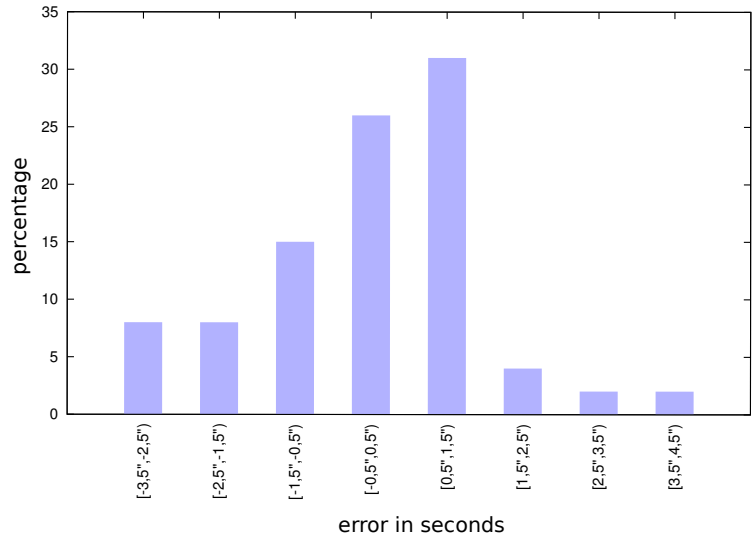


Figure 4.14: Difference in seconds between the expert average and the automatic BUT measurements.

4.6 Analysis of Break-up Classification

The break-up classification is performed in the ERF of each SOI, since it is the point where the break-up reaches its maximum expansion. Therefore, in these frames, the break-up is segmented, and then, the contour of each break-up area is extracted. A descriptor for each contour is created from its features, this is, area (A), ratio between and major and minor axes (r_{ab}), roundness (κ), eccentricity (ϵ), convexity (C), and solidity (S). The following classifiers are used over these descriptors in order to classify each break-up area in one of the rupture patterns:

- **Naive Bayes** [Jensen, 1996] is a simple probabilistic classifier based on the Bayesian theorem which can predict class membership probabilities.
- **Support Vector Machine (SVM)** [Burges, 1998] is based on the concept of decision hyperplanes that separate all member of one class from those of the other class.

- **Multilayer perceptron (MLP)** [Rosenblatt, 1958] is a feedforwad artificial neural network which consists of a set of layers of nodes, joined together in a pattern of connections.
- **Decision Tree** [Rokach and Maimon, 2006] is a logic-based algorithm which classifies samples by sorting them based on feature values.

The break-up classification is validated in a dataset of 88 tear film frames containing 25 dots, 33 pools and 38 streaks rupture patterns annotated by 2 experts. Due to the limited size of our data set, 10-fold cross-validation [Rodriguez et al., 2010] was used for assessing the generalization capability. Table 4.8 shows the performance by each classifier for the rupture patterns classified by the automatic system in relation to the experts' annotations.

Table 4.8: Performance of break-up classification by comparing the automatic and manual annotations.

	System			
	Dots	Pools	Streaks	
Bayes	92.00	4.00	4.00	Dots
	9.09	63.63	27.27	Pools
	2.63	0.00	97.36	Streaks
SVM	68.00	12.00	20.00	Dots
	0.00	69.70	30.30	Pools
	2.63	5.26	92.10	Streaks
MLP	80.00	16.00	4.00	Dots
	9.09	66.67	24.24	Pools
	5.26	13.16	81.58	Streaks
DTree	92.00	4.00	4.00	Dots
	3.03	81.82	15.15	Pools
	5.26	13.16	81.57	Streaks

Expert

The decision tree classifier got the best results achieving 84.37% global accuracy for all rupture patterns. In some cases, very small pools were confused with dots as well as dots quite large were classified as pools. Some poorly defined streaks were classified as pools while streaks too small where mistaken with dots.

In order to verify the independence of the features in the descriptor and to extract the most relevant ones, principal component analysis (PCA) [Jolliffe, 1986] is applied to the decision tree classifier. The extracted features by applying PCA are the axis ratio, the roundness, the eccentricity, the area, and the solidity. These features cover the main discriminant factors between the different patterns. The convexity is discarded by PCA because it provides similar information to the solidity since both are computed from the convex hull area, which may be redundant. The decision tree applied after discarding this feature achieved the results showed in Table 4.9. The dot and streak patterns have increased their sensitivity whereas the sensitivity of pool pattern has decreased. The global accuracy was 85.42%, which represents a slight increment respect to results obtained from the whole descriptor.

Table 4.9: PCA performance of break-up classification by comparing the automatic and manual annotations.

	System				
	Dots	Pools	Streaks		
	96.00	0.00	4.00	Dots	
PCA	6.06	75.76	18.19	Pools	**Expert**
	5.26	7.89	86.84	Streaks	

4.7 Evaluation of the Local Analysis

The tear film presents different conditions based on the location due to eye features. Therefore, the central cornea is much more sensitive than the periphery, so tear film break-up at this zone should affect vision as a result of tear film changes over the pupil. Moreover, the break-up probability is not the same at each zone since the tear film is more stable in the lower part of the eye. In order to extend the global methodology to the local analysis, the ROI is divided into five equally sized zones according to the Cornea and Contact Lens Research Unit (CCLRU) standards, and each of the break-up

features evaluated in the global analysis is performed separately in each local zone. This analysis provides a more specific break-up characterization since it is focused on local eye features.

To this end, first a specific break-up threshold is computed for each local zone. For this purpose, multilevel thresholding is performed over each local zone of the ERF, since this approach achieved the best results in global segmentation. Thus, $4-level$, $6-level$, $8-level$, and $10-level$ thresholding are analyzed for the different k-fold training sets in each local zone, and then, each threshold t_b is computed from the upper limit of l_b, this is, the level which provides better results at that zone. The different 10 k-fold training sets provide the same results for the best l_b in each local thresholding. Thus, mean ROC curves for each local zone computed from the mean values of sensitivity and specificity in the 10 k-fold training sets are used as a representation of the training iteration, as shown in Fig. 4.15.

The $4-level$ was discarded since it presents results significantly worse at the zone 4. The remaining thresholding configurations provides similar results, being the $8-level$, and $10-level$ slightly better. This way, the $8-level$ thresholding is selected for all the local zones, and l_b is set to levels 3, 3, 4, 3, and 4 for the local zones from 0 to 4, respectively, since they provide the best compromise between sensitivity and specificity. These values were checked in the 10 k-fold test sets providing the accuracy results showed in Table 4.10. The local segmentation with these values achieves acceptable results considering the high variability between the experts. Appendix C.3 includes the ROC curves for each k-fold training set and the accuracy results of each k-fold test set.

Once the specific break-up thresholds are set, an evolution curve for each local zone is build from its corresponding threshold. Then, other specific thresholds are computed from a percentage of the maximum height of each local curve to get the BUT measurement. Thus, a local BUT measurement

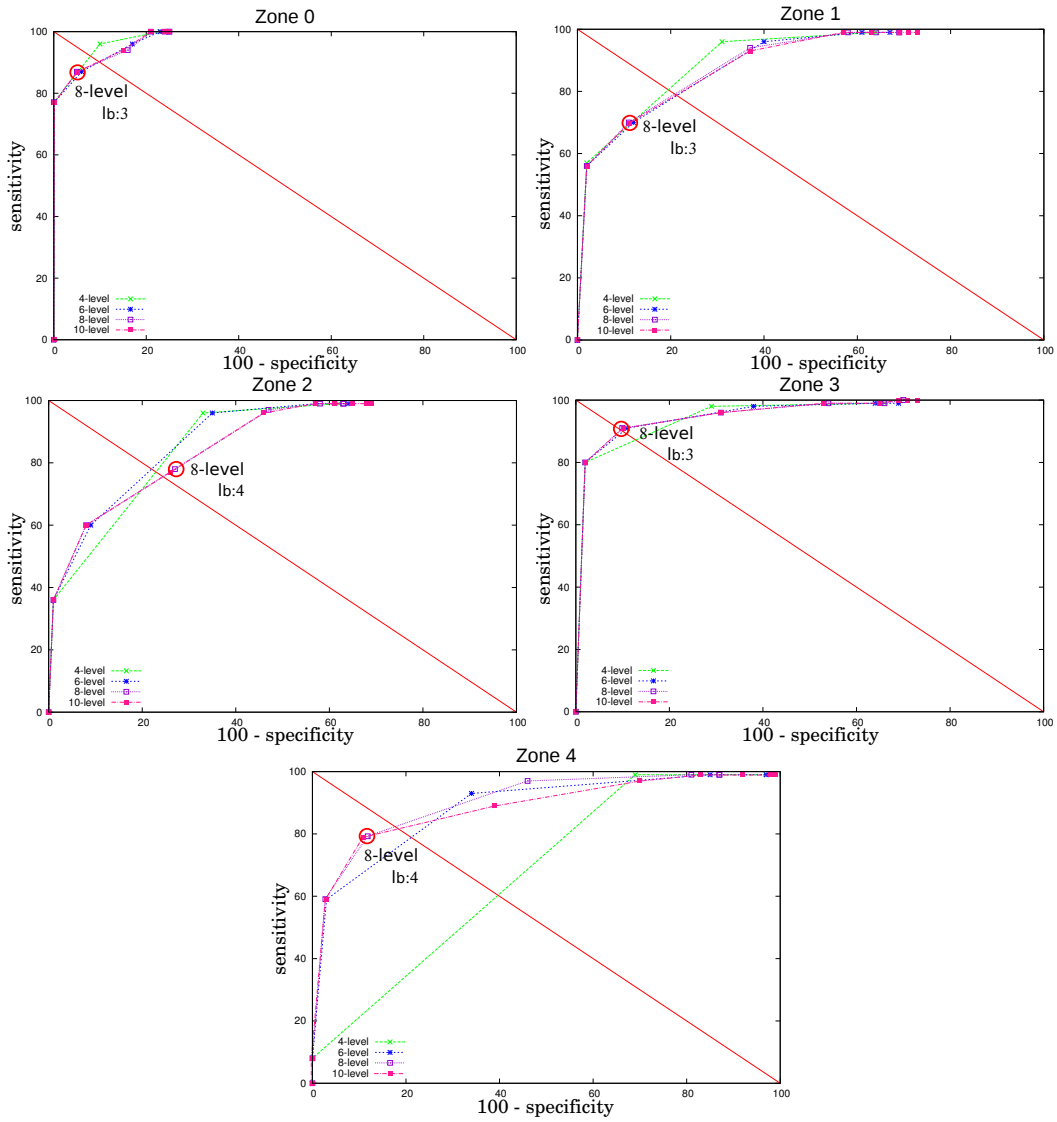


Figure 4.15: Mean ROC curves of each local zone for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

Table 4.10: Accuracy (%) of break-up segmentation at each local zone for the best case, the worst case, and the mean of the 10-kfold test sets.

	Global ($6-level$ $l_b=3$)	Zone 0 ($8-level$ $l_b=3$)	Zone 1 ($8-level$ $l_b=3$)	Zone 2 ($8-level$ $l_b=4$)	Zone 3 ($8-level$ $l_b=3$)	Zone 4 ($8-level$ $l_b=4$)
Best case	92.84	97.72	91.12	88.41	93.24	89.16
Worst case	86.87	91.90	72.58	74.85	86.58	74.22
Mean	89.09	93.55	83.64	83.08	90.25	83.57

is obtained for each local zone, and finally, the global BUT measurement is obtained from the earliest local BUT, since it represents the first appearance of dark areas, that is, the tear film break-up.

The ROC curve is built with p_e taking values from 5% to 95% of the total height of the evolution curve for each k-fold training set. The best value for p_e is always 95% in the different k-fold iterations, so the ROC curve computed from the mean values of specificity and sensitivity in the 10 k-fold training sets is used to set the value of p_e, as shown in Fig. 4.16. Thus, p_e is set to 95%, since it provides the best compromise between sensitivity and specificity. This value was checked in the 10 k-fold test sets providing the results showed in the Table 4.11. Appendix C.4 includes the ROC curves for each k-fold training set as well as the accuracy results of each k-fold test set.

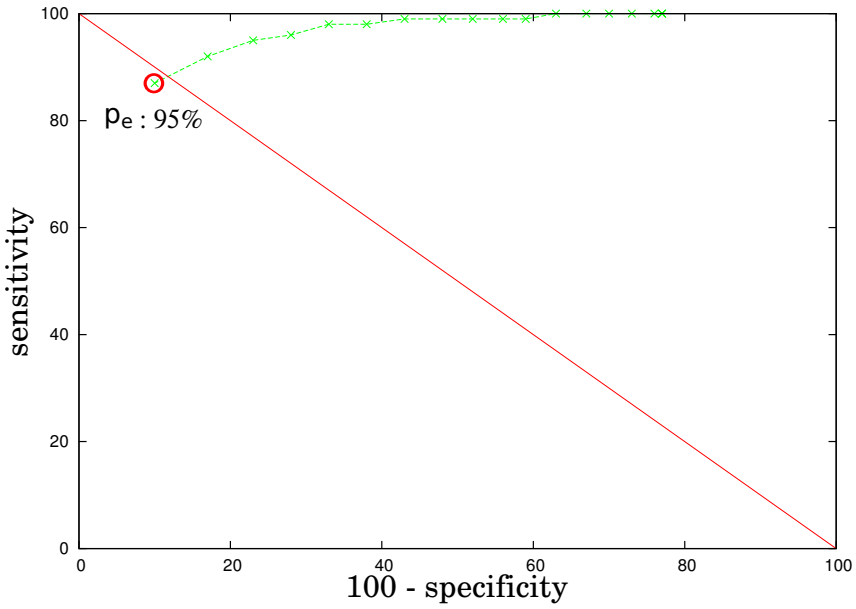


Figure 4.16: Mean ROC curve for the 10 k-fold training sets obtained with percentages from 5% to 95% of the total height of the local evolution curves. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

Table 4.11: Accuracy (%) of BUT measurement for the best case, the worst case, and the mean of the 10 k-fold test sets with the best value for p_e in the global and local analysis.

	Global ($p_e = 65\%$)	Local ($p_e = 95\%$)
Best case	90.43	94.14
Worst case	86.22	84.36
Mean	86.73	90.14

After that, the automatic BUT measurements obtained from the analysis of local evolution curves have been compared to the experts' average in the full dataset, as shown in Fig. 4.17. As can be seen, the local version improves the BUT measures obtained with the global version. Concretely, the percentage of values within the range of ±2.5 seconds is 95.55%, which is practically the same percentage in the same range presented among the experts themselves. Moreover, the mean deviation was reduced to 0.96 seconds and the percentage of values below ±0.5 seconds increased over 11%. In contrast to global measurement, local approach tends to anticipate the BUT since it is more sensitive to intensity variations at each local zone.

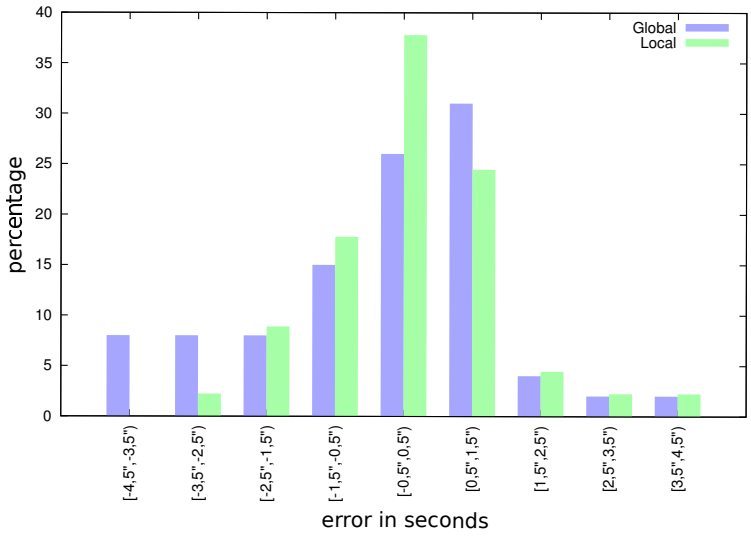


Figure 4.17: Dispersion between the expert average and the BUT measurement for the global and local versions.

Considering each local zone presents different features and break-up probabilities, an additional analysis is performed over each evolution curve. To this end, a local BUT measurement is computed from each evolution curve and compared to the results provided by the experts. This way, optimum height percentage for each curve is computed. The ROC curves are built with p_e taking values from 5% to 95% of the total height of each local evolution curve for each k-fold training set. As in the previous cases, the ROC curves computed from the mean values of specificity and sensitivity in the 10 k-fold training sets are used to set the values of each local p_e, as shown in Fig. 4.18. Thus, the local p_e are set to 65%, 90%, 70%, 55% and 75% of the maximum heights of the evolution curves from zones 0 to zone 4, respectively. These values were checked in the 10 k-fold test sets providing the accuracy results showed in Table 4.12. Appendix C.3 includes the ROC curves for each k-fold training set as well as the accuracy results of each k-fold test set.

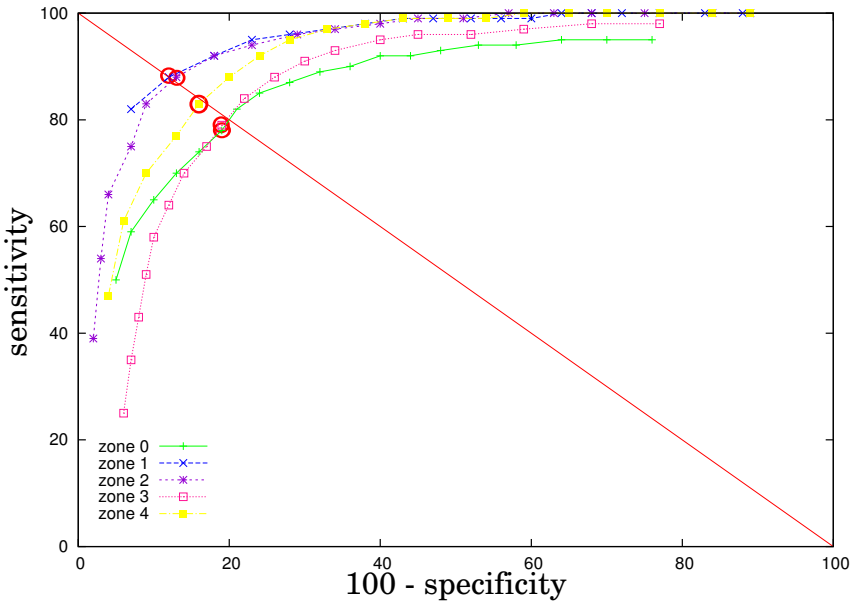


Figure 4.18: Mean ROC curves for the 10 k-fold training sets obtained with percentages from 5% to 95% of the total height of the local evolution curves. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements obtained from each local evolution curve.

Table 4.12: Accuracy (%) of break-up segmentation for the best case, the worst case, and the mean of the 10 k-fold test sets with the best values for p_e.

	Local ($p_e = 95\%$)	Zone 0 ($p_e = 65\%$)	Zone 1 ($p_e = 90\%$)	Zone 2 ($p_e = 70\%$)	Zone 3 ($p_e = 55\%$)	Zone 4 ($p_e = 75\%$)
Best case	94.14	75.96	84.11	80.09	74.65	77.92
Worst case	84.36	86.74	93.21	94.84	87.55	90.09
Mean	90.14	80.06	90.02	86.42	80.15	83.37

Once a specific p_e is set for each local zone, the BUT measurements obtained from the first local BUT with this parameter configuration are compared to the experts' average in the full dataset. Figure 4.19 shows the dispersion between the manual values and the automatic measurements from the local approaches using the same and different thresholds for each local zone. The percentage of values below ± 2.5 seconds obtained with these thresholds is 88.89% and the mean deviation is 1.24 seconds. This value is slightly higher than in the previous approach but is in the same range as among the experts themselves. This version tends to get ahead of the expert measure since it is focused at each local zone and its specific features.

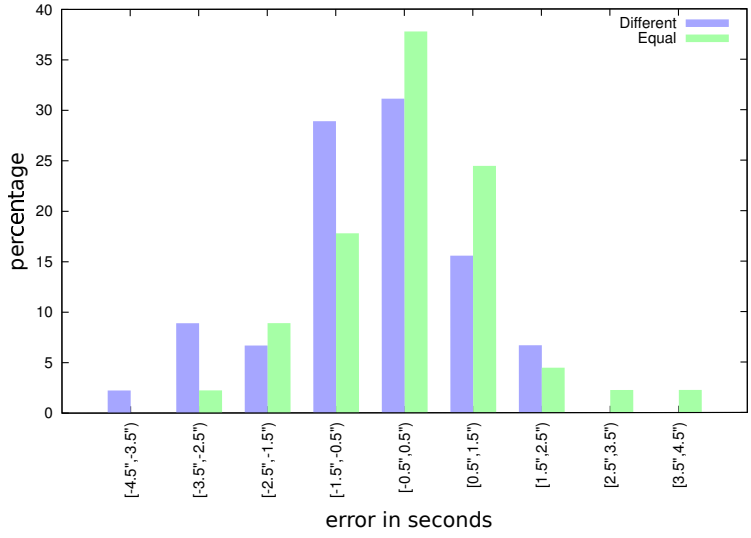


Figure 4.19: Dispersion between the expert average and the BUT measurement for the local analysis using the same and different thresholds for the local zones.

4.8 Performance

Each tear film video has a duration of several minutes and contains an average of four different SOIs which are analyzed at different execution threads. The execution time of each entire video is about 1.5 minutes, in which the location of the SOIs takes about 11 seconds and the break-up assessment of each of them lasts about 45 seconds. Table 4.13 shows the times for the different steps performed for analyzing each SOI. Therefore, this method can approximate to real time use providing results that can be applied in the clinical practice.

Table 4.13: Time performance (seconds) of the different steps for break-up assessment.

Step	Time
SOI Location	10.598
ROI Extraction	26.547
Normalization	1.191
Break-up Segmentation	0.018
Break-up Time	3.380
Pattern Classification	0.087

Chapter 5

Conclusions and Future Work

Dry eye syndrome (DES) is a common clinical condition which affects the general population, impacting on quality of life. The discomfort and visual disturbance associated with dry eye syndrome is a growing public health concern. The diagnosis of this disease is a difficult task due to its multifactorial etiology, and so there exist several clinical tests to evaluate different aspects of the tear film. One of these tests is the BUT test, which consists in measuring the time that the tear film remains stable without blinking. To perform this test, sodium fluorescein is instilled into the eye, and the BUT is measured as the time elapsed until a dark area appears in the tear film, which represents the evaporation of water and the break-up of the tear film. A low break-up measurement correspond to a limited ocular surface wetting, and it is one of the main signs of an abnormal tear film.

This test only examines the appearance of the first dark spot in the tear film, regardless the subsequent break-up dynamics. However, there are other break-up properties related to specific aspects of the tear film that could affect to dry eye severity. On one hand, the break-up location is a relevant feature since there are corneal zones specially vulnerable to tear disruptions. On the other hand, the first break-up could appear as a small point or as a wide area

and it could increase its size with time presenting different rupture patterns, which are related to the BUT measure, the break-up area, and its growing rate. This information is omitted in the original BUT test, but could be relevant in tear film assessment for understanding the tear film instability in DES and its relation to ocular surface symptoms.

BUT test is affected by low repeatability mainly due to a subjective appreciation of the dark spots, the differences among the experts, and the variability of the tear film. Besides the subjectivity, the characterization by hand of break-up dynamics is a tedious and time consuming task. The automation of the break-up assessment would reduce its subjective character, allowing a more accurate evaluation of the tear film stability. In this work, an automatic methodology for assessing tear film break-up over the exposed corneal surface has been developed in order to evaluate the quality and stability of the tear film, essential for DES characterization. This analysis includes the BUT measurement computation as well as a quantitative and qualitative analysis in terms of location, size, shape and dynamics of break-up areas.

Each tear film video is divided in several BUT sequences, denoted as Sequences of Interest (SOIs), which consists of a set of frames delimited by blinks where the BUT test can be performed. The Region of Interest (ROI) for break-up analysis corresponds to the visible part of the iris which may vary slightly throughout the sequence depending on the eye aperture and the appearance of shadows due to outer parts of the eye like eyelids or eyelashes.

In the first stage, the tear film videos have been preprocessed to extract the SOIs in the sequence and after that, the ROI in each frame. The SOIs have been delimited by detecting the blinks from symmetric finite differences of mean values of gray between consecutive frames. After that, the ROI within each frame has been extracted by correlation in the frequency domain with a set of elliptical masks. Then, the ROI has been registered throughout the

SOI by aligning the ROI at each frame with the immediately preceding frame. Thus, the methodology is independent of slight motions of the eye since the ROI appears always in the same location. Moreover, an adaptive adjustment according to the eye features has been performed to discard outer parts like eyelids or eyelashes. For this adjustment, the mask radius has been slightly reduced and the ROI has been cropped at the top and at the bottom. The SOI and ROI extraction have performed quite well, providing a well defined space for break-up analysis, with an accuracy over 90%.

Once the SOIs have been delimited and the ROI has been extracted within each frame, the key stage of the methodology is the break-up assessment. It consists of the analysis of the different SOIs in order to detect the break-up time and characterize the break-up areas. For this purpose, first, a preprocessing step has been performed to normalize the lightness and contrast variability in tear film frames and then, a break-up threshold has been set to determine the break-up level in each SOI, that is, the maximum intensity for a pixel to be considered as part of the break-up. For this purpose, two alternatives have been explored. In the first approach, the break-up threshold has been computed from a percentage of the darkest pixels of the cumulative histogram of the Start Reference Frame (SRF). In the second one, the break-up threshold has been obtained from one of the lowest levels resulting by applying multilevel thresholding to the End Reference Frame(ERF). The break-up segmentation has achieved accuracy results around 85% in the first approach and 90% in the second one, which are acceptable results considering the coincidence between experts is around 60%.

After the preprocessing step, the break-up assessment has been performed in each SOI. For this purpose, the BUT measurement has been computed first, as the time elapsed from the beginning of the sequence until the emergence of the first break-up in the tear film, that is, the appearance of dark areas on

the surface of the eye. In order to detect the emergence of these points, the percentage of break-up pixels has been computed for each frame, and then, the evolution of these percentages has been examined through the SOI. The BUT measurement hs been computed as the time elapsed from the beginning of the SOI until the evolution curve exceeds a percentage of its maximum height. The variability among the experts is in an interval of ± 2.5 seconds due to the subjectivity of this measurement. The global BUT measurement has got accuracy results around 88%, being the mean deviation 1.21 seconds. This value is in agreement with the manual detection, since it is within the same range as among the experts themselves.

The break-up areas could increase its size with time presenting different rupture patterns. Therefore, after the computation of the first BUT, the break-up areas have been analyzed until the ERF. On one hand, the methodology has included an step in order to classify the segmented break-up areas into the different rupture patterns. To this end, a shape descriptor containing morphological features such as roundness, eccentricity, ratio between the major and minor axes, area, convexity, and solidity has been built for each segmented break-up area, and then, different classifiers have been applied to these descriptors to decide their tear film rupture patterns. The classification has achieved a global accuracy over 85%. Furthermore, the size and growing rate of the break-up areas have been measured for characterizing the break-up dynamics. This characterization allows to distinguish between an early or late break-up, and if the first break-up appears as a small point or as a wide area and if it increases fast or slow.

The break-up location is related to specific aspects of the tear film that could affect to dry eye severity. In order to extend the global methodology to the local analysis, the ROI has been divided into five equally sized zones according to the Cornea and Contact Lens Research Unit (CCLRU) stan-

dards, and each of the break-up features considered in the global analysis has been evaluated separately in each local zone, obtaining a more specific break-up characterization. The local break-up segmentation has achieved accuracy results between 83% and 93% in the different local zones. The local BUT measurement has improved the results obtained with the global version. Concretely, the percentage of values within the range of ± 2.5 seconds has been 95.55%, which is practically the same percentage in the same range presented among the experts themselves. Moreover, the mean deviation has been reduced to 0.96 seconds. Furthermore, the local evolution curves have more quality since they are representing specific areas focused on the break-up zone.

Summarizing, a novel methodology has been developed in this work for a fully automatic assessment of the tear film break-up. The proposed methodology allows a quantitative, qualitative analysis of tear film instability, as an extension of BUT measurement, which is focused only on time. It provides accuracy results in the same ranges as the experts. The results of this work prove that the automation of break-up assessment is not only feasible but also offers new opportunities for diagnosis and screening. The automated system provides fast, objective results, saving on effort, time and costs. It provides the possibility of large population studies to extract new variables of interest for clinical practice.

5.1 Further work

Future work in this field includes an analysis of the repeatability of the BUT test and break-up dynamics between different SOIs of the same patient to evaluate if certain tendencies are repeated in order to detect weakened areas of the ocular surface. Many physiological variables oscillate during the day, such as intra-ocular pressure corneal sensitivity or factors that affect tear quality and

dry eye symptoms. The possibility of diurnal variations in tear film variables should be taken into account by the clinicians, so an analysis of the variability among tear film measurements captured at different time of the day could be interesting.

The break-up assessment could be combined with other clinical tests for analyzing the relationship between them. Since in daily clinical practice it is often used a short protocol with some common tests for the study of the anterior segment of the eye, their results could be analyzed and correlated for extracting new useful variables to get a dry eye diagnosis based on different factors.

Furthermore, the data set could be extended to include more representative cases with more patients and for a longer follow-up period to analyze the evolution among different clinical checkups.

Appendix A

Acronyms

DES Dry Eye Syndrome 4

BUT Break-Up Time 7

ROI Region Of Interest 29

SOI Sequence Of Interest 27

SRF Start Reference Frame 36

ERF End Reference Frame 36

TP True Positives 78

TN True Negatives 78

FP False Positives 78

FN False Negatives 78

Appendix B

Adapted Canny Edge Detector

The ROI extraction step requires the edge image of the tear film frames for segmenting the iris. Assuming the iris has an approximately circular shape, its initial segmentation is carried out by correlation of the edge image with a set of masks formed by circumferences with different radii. This way, a suitable edge representation for this step would contain all possible pixels related to iris boundaries and would avoid another edge pixels which could mismatch with the circular mask. For this purpose, an adaptation of Canny edge detector [Canny, 1986] is implemented.

B.1 Original Algorithm

The original Canny edge operator used for detecting the edges in the ROI extraction is a multi step algorithm based on three main criteria:

- Maximize the signal/noise ratio to favor the detection of true positives.
- Minimize the distance between edge pixels detected and real edge pixels.
- Minimize the number of responses to an edge, to reduce the number of false positives .

This is composed of the following steps:

- **Noise reduction** A smoothing of the frame is performed by convolution with a Gaussian filter in order to remove noise. The result is a slightly blurred version of the original frame which is not affected by a single noisy pixel to any significant degree.

- **Derivation** The edges are related to pronounced intensity variations, which are given by the gradient of the image. Thus, a first approximation of the edges is computed from the first derivative in the x and y by applying the following differential filters:

$$G_x = \begin{vmatrix} -1 & 0 & +1 \end{vmatrix} \qquad G_y = \begin{vmatrix} -1 \\ 0 \\ +1 \end{vmatrix} \tag{B.3}$$

  Then the magnitude and direction of the gradient are computed as follows:

$$\begin{aligned} G &= \sqrt{G_x + G_y} \\ \theta &= \arctan\left(\frac{G_y}{G_x}\right) \end{aligned} \tag{B.4}$$

- **Non-maximal suppression** The edge magnitude could contain weak edges around de local maximum due to multiple response. The aim of this step is to discard these pixels in order to get one pixel thick edges. For this purpose, only the local maximum of the magnitude in the ortogonal gradient direction are selected.

- **Hysteresis** The final step is used to remove the weak edges remaining after the non-maximal suppression. To this end, a hysteresis function is applied in order to identify the points that actually belong to an edge using an upper and a lower threshold. Thus, pixels with gradient higher than the upper threshold are accepted as edges, while pixels with gradient

below the lower threshold are rejected. The pixels between the two limits are accepted only if there is a path to join them to an actual edge.

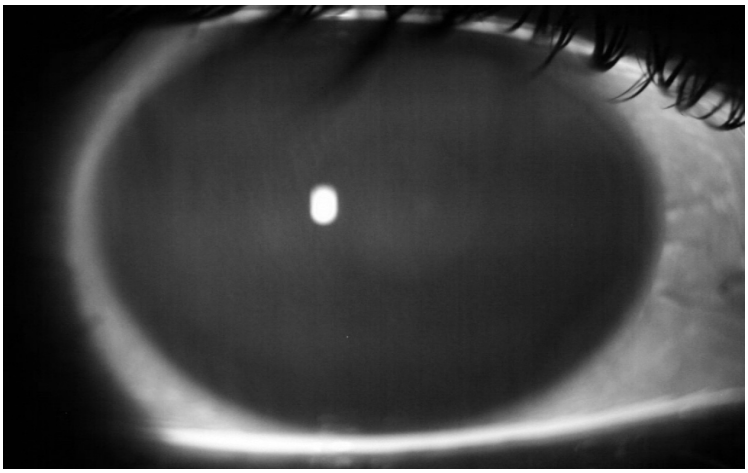
B.2 Adaptation for Iris Segmentation

The Canny edge operator described in the previous section is applied to the green component of the tear film frame. The different parameters used by the algorithm are set empirically. This way, the size of the Gaussian filter σ is set to 9, and the higher and lower thresholds used in the hysteresis step are set to 0.35 and 0.55, respectively.

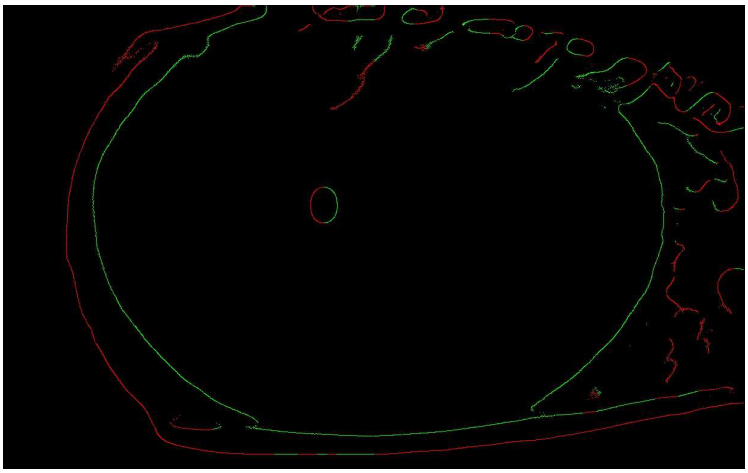
The natural structure of the eye contains the sclera, a white fibrous membrane that forms a circular region surrounding the iris (see Fig. B.1 (a)). The edge pixels related to this area could disrupt the ROI extraction, so they are discarded. To this end, the orientation of the gradient computed by the Canny edge detector is considered. The gradient orientation is computed as the angle of a vector centered at the edge pixel and pointing towards the brightest area in the direction of maximum intensity change. This way, the edge pixels corresponding to the iris have outgoing orientation in relation to the iris center since they are surrounded by the bright part related to the sclera. However, the sclera is surrounded by a darker area, so its edge pixels have incoming orientation (see Fig. B.1 (b)). Thus, an edge image E' (see Fig. B.1 (c)) is computed by discarding the incoming edge pixels in relation with the center (c_i, c_j) of the original edge image E as follows:

$$E'(i,j) = \begin{cases} 1 & if \quad E(i,j) = 1 \wedge j \geq c_j \wedge \theta(i,j)\epsilon(\frac{\pi}{2}, \frac{3\pi}{2}) \\ 1 & if \quad E(i,j) = 1 \wedge j < c_j \wedge \theta(i,j)\epsilon(\frac{3\pi}{2}, \frac{\pi}{2}) \\ 0 & otherwise \end{cases} \tag{B.5}$$

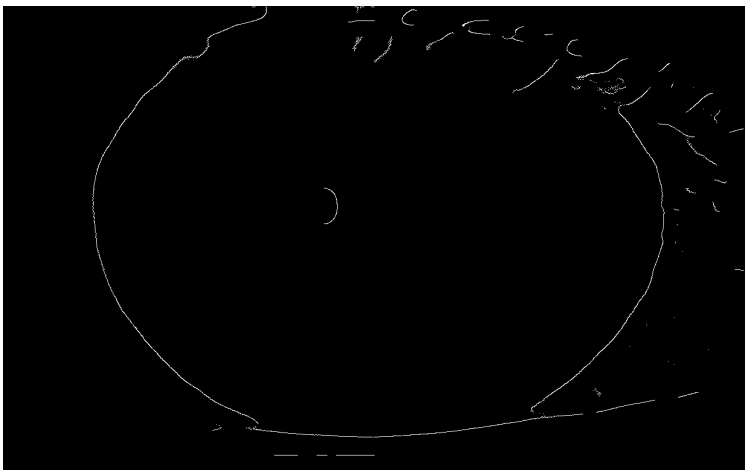
where θ is the orientation of the gradient from Canny edge detector.



(a)



(b)



(c)

Figure B.1: Discarding incoming pixel from Canny edge detection. (a) Original tear film frame. (b) Edge image E obtained by applying Canny. Green pixels are related to outgoing edges whereas red pixels represents the incoming edges in relation to the center. (c) Edge image E' after discarding incoming pixels.

Appendix C

k-fold Validation

Due to the limited size of the dataset, 10-fold cross-validation is used in the experiments to assess the generalization capability [Rodriguez et al., 2010]. Therefore, the dataset is divided into 10 parts and an iterative process is carried out 10 times. At each iteration, one of the parts is used as a test set, and the remaining parts are used as a training set. Finally, ROC curves are built for each 10 k-fold training set and they are analyzed to adjust each parameter. Once the best parameter values are selected, the accuracy is computed from the k-fold test sets.

C.1 Global Segmentation

The break-up segmentation is validated considering that each pixel can be classified as a break-up pixel or a background pixel. This way, values for sensitivity and specificity are obtained by comparing the matches pixel by pixel between the manual break-up regions marked by the experts and the automatic results provided by the system.

Figure C.1 shows the ROC curves built with p_b taking values from 5% to 95% for the 10 k-fold training sets. The best values for p_b are always between

40% and 45% of the darkest pixels in the cumulative histogram of the SRF. Figure C.2 shows the ROC curves built with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding of the ERF. The $6-level$ presents slightly better results than the $4-level$, $8-level$, and $10-level$ thresholding. The best value for l_b in $6-level$ is set to 3 in the 10 k-fold training sets. Once the best values are set, their accuracy is computed in the 10 k-fold test sets. Table C.1 contains the accuracy results of each k-fold test set. Both approaches get acceptable results considering the high variability between the experts.

Table C.1: Accuracy (%) of break-up segmentation for the 10 k-fold test sets with the best values for p_b and l_b.

	Cumulative histogram ($p_b = 45\%$)	Multilevel thresholding ($6-level, l_b = 3$)
k1	90.94	90.61
k2	86.59	91.87
k3	84.84	92.84
k4	83.32	87.55
k5	87.41	88.73
k6	85.76	89.06
k7	77.33	88.19
k8	82.69	87.71
k9	86.98	86.87
k10	83.30	87.42
Mean	84.92	89.09

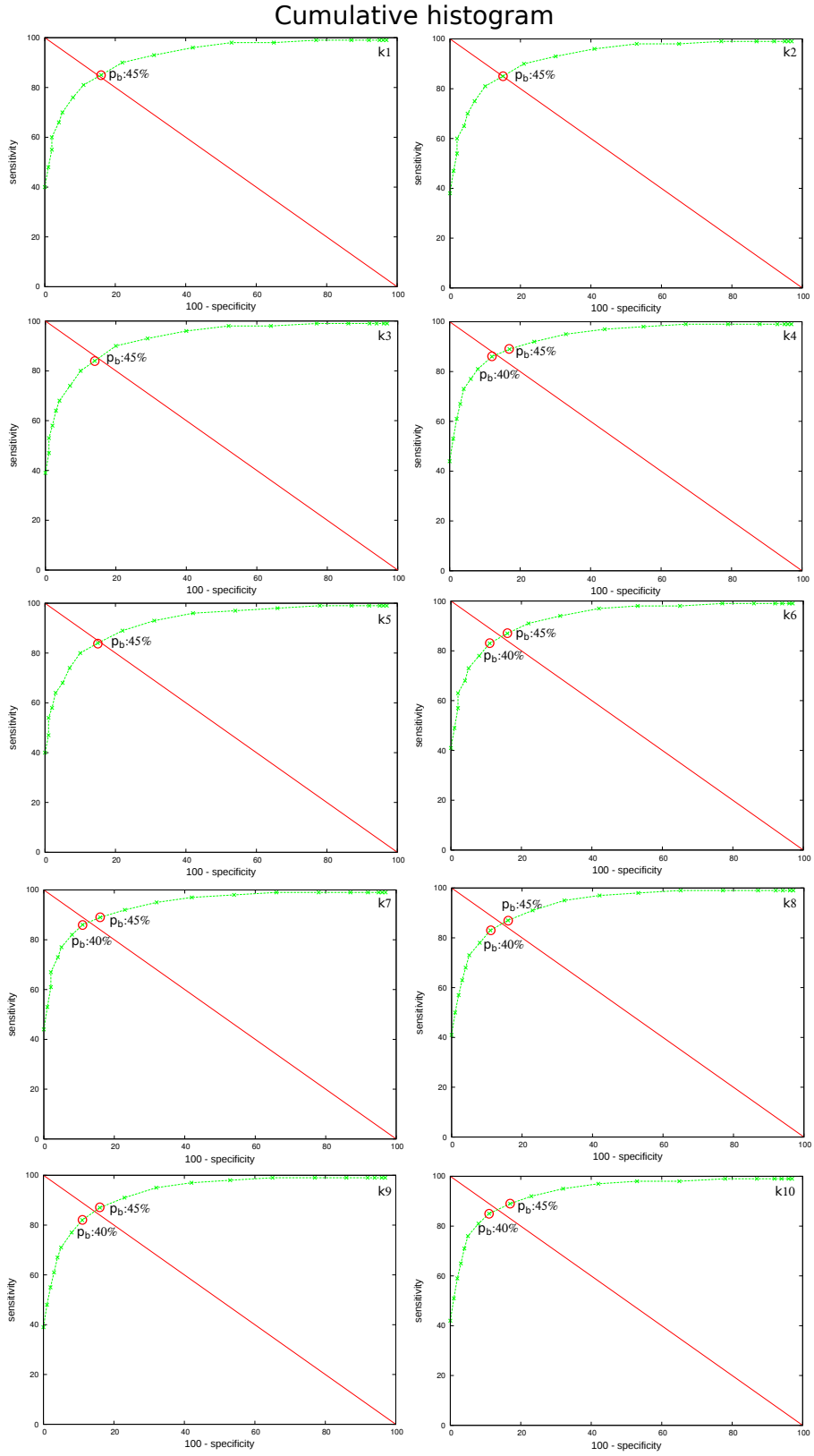


Figure C.1: ROC curves for the 10 k-fold training sets obtained with percentages p_b from 5% to 95% of the darkest pixels in the cumulative histogram of the SRF. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

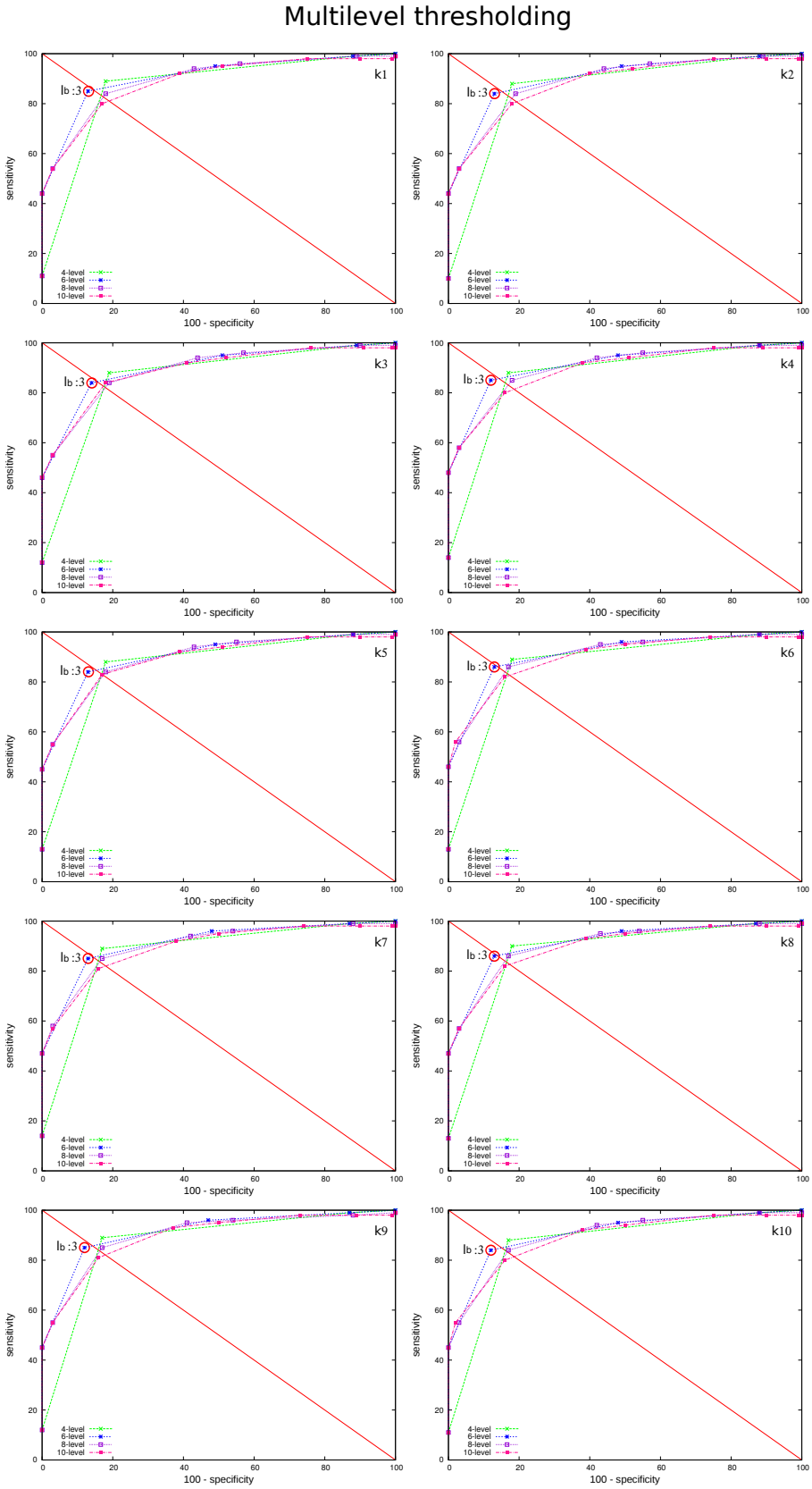


Figure C.2: ROC curves for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding of the ERF. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

C.2 Global BUT

The BUT measurement is validated considering that each frame can be classified as a frame with or without break-up. This way, the frames located before the BUT are considered as frames without break-up. Similarly, the subsequent frames to the BUT are considered as frames with break-up. Therefore, values for sensitivity and specificity are computed from the coincidences between manual and automatic measurements of the frames classified as break-up frames and those classified as frames without break-up,

Figure C.3 shows the ROC curves built with p_e taking values from 5% to 95% for the 10 k-fold training sets. The best values for p_b are always between 60% and 65% of the total height of the evolution curve. Once the best value is set, its accuracy is computed in the 10 k-fold test sets. Table C.2 contains the accuracy results of each k-fold test set.

Table C.2: Accuracy (%) of BUT measurement for the 10 k-fold test sets with the best value for p_e.

	Global ($p_e = 65\%$)
k1	90.43
k2	86.22
k3	80.13
k4	81.94
k5	92.17
k6	90.33
k7	83.55
k8	88.66
k9	83.94
k10	89.96
Mean	86.73

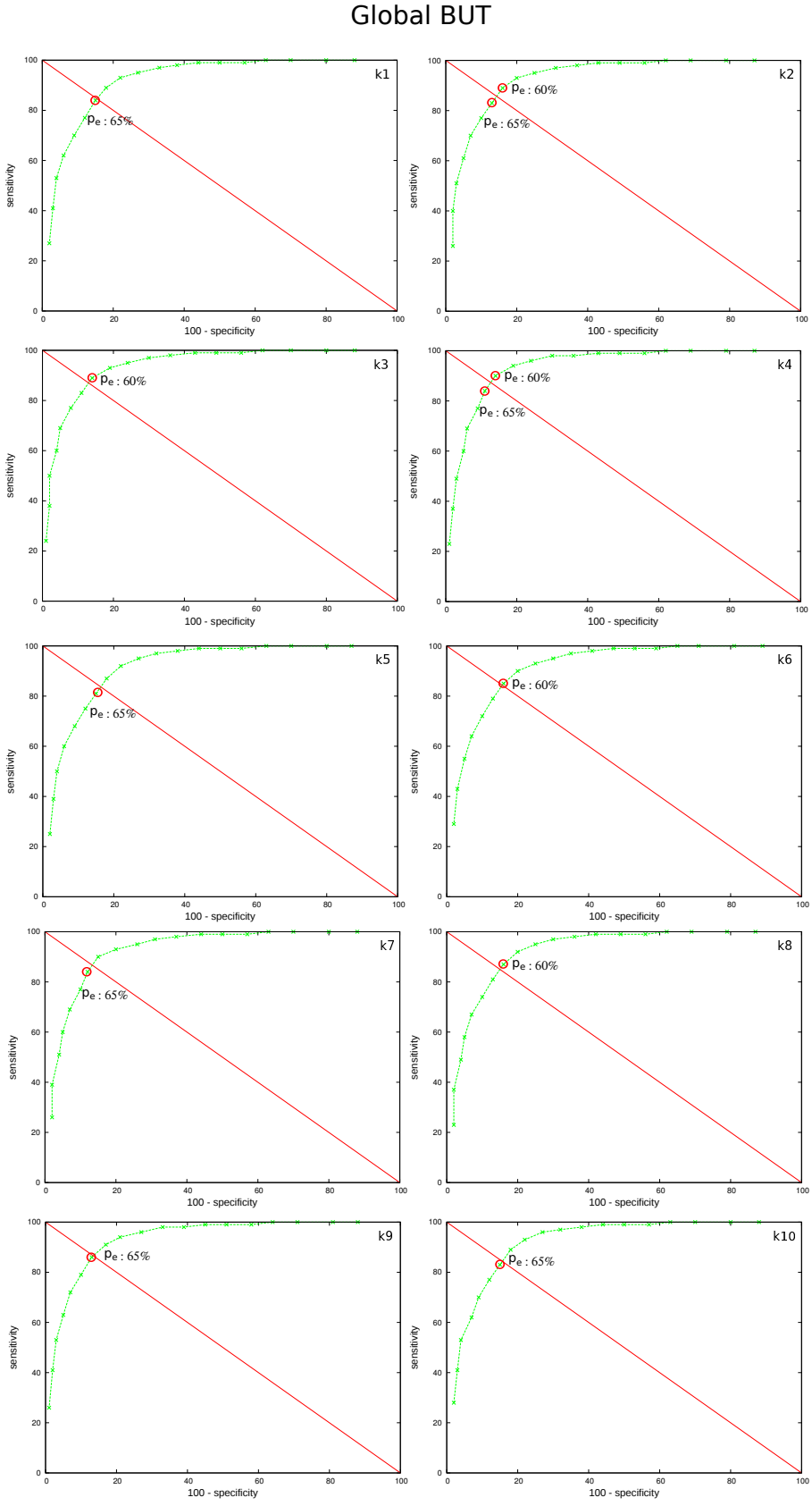


Figure C.3: ROC curves for the 10 k-fold training sets obtained with percentages p_e from 5% to 95% of the maximum height of the evolution curve. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

C.3 Local segmentation

In the local analysis, a specific break-up threshold is computed for each local zone. The validation of local break-up segmentation is performed by comparing the matches pixel by pixel between the manual break-up regions marked by the experts and the automatic results provided by the system at each local zone.

Figures C.4, C.5, C.6, C.7, and C.8 show the ROC curves for the 10 k-fold training sets obtained from $4-level$, $6-level$, $8-level$, and $10-level$ thresholding at zones from 0 to 4 of ERF, respectively. The different 10 k-fold training sets provide the same results for the best l_b in each local thresholding. This way, mean ROC curves for each local zone represent the training iteration in the 10 k-fold training sets. Once the best value are set for the local zone, their accuracy is computed in the 10 k-fold test sets. Table C.3 contains the accuracy results of each k-fold test set.

Table C.3: Accuracy (%) of break-up segmentation for the 10 k-fold test sets with the best values for l_b at each local zone.

	Zone 0 ($8-level$ $l_b=3$)	Zone 1 ($8-level$ $l_b=3$)	Zone 2 ($8-level$ $l_b=4$)	Zone 3 ($8-level$ $l_b=3$)	Zone 4 ($8-level$ $l_b=4$)
k1	97.72	86.19	86.16	93.24	84.62
k2	92.51	82.86	83.88	90.66	86.73
k3	91.99	89.23	81.63	92.13	89.16
k4	97.31	72.15	84.26	86.58	74.22
k5	88.08	91.12	82.61	86.05	85.49
k6	92.92	88.04	74.85	91.81	83.98
k7	91.90	72.58	88.41	87.57	77.76
k8	91.98	86.18	82.86	90.07	77.99
k9	93.64	90.95	81.77	91.54	89.00
k10	97.48	77.05	84.37	92.84	86.73
Mean	93.55	83.64	83.08	90.25	83.57

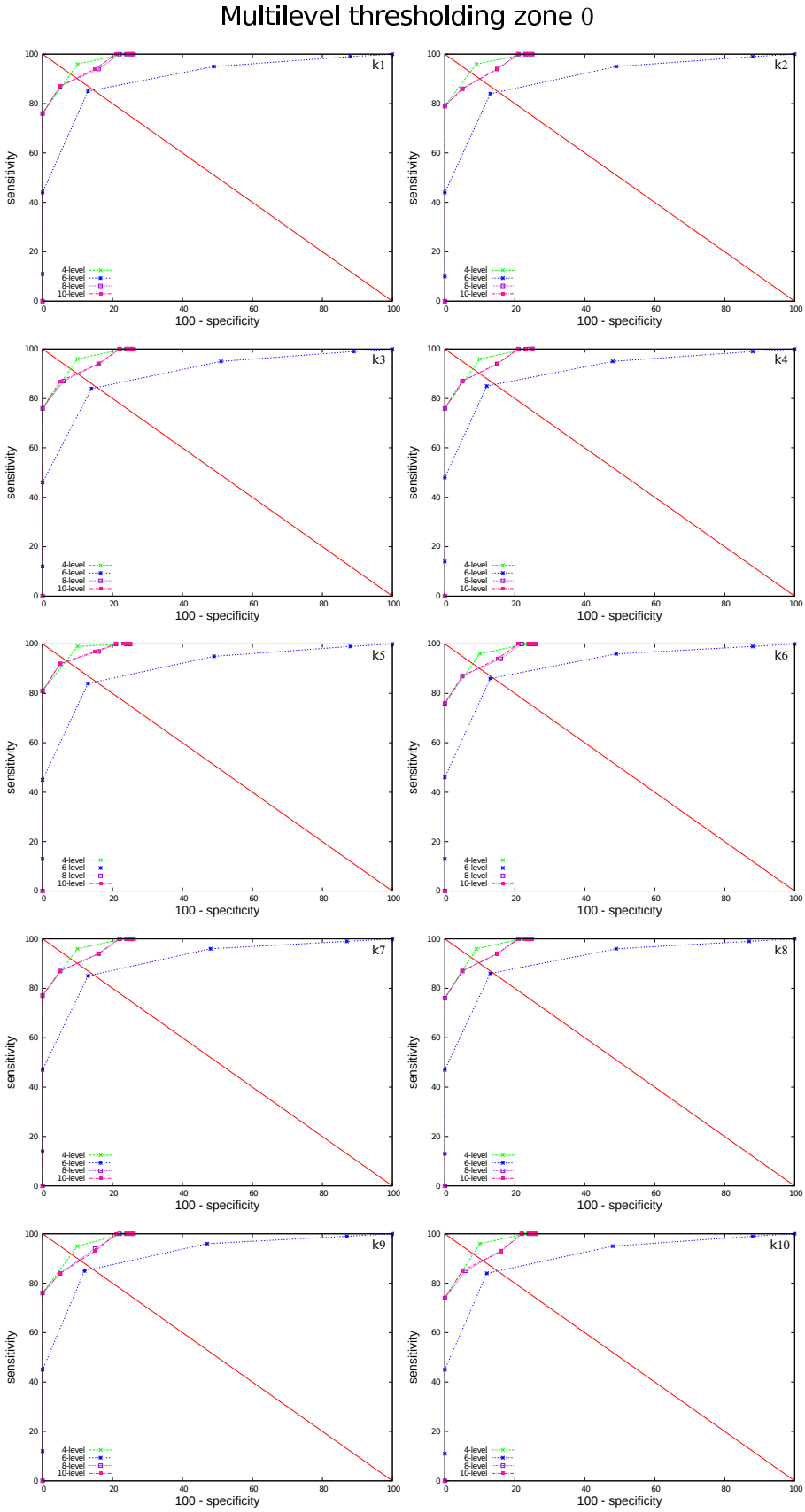


Figure C.4: ROC curves for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding at zone 0 of ERF. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

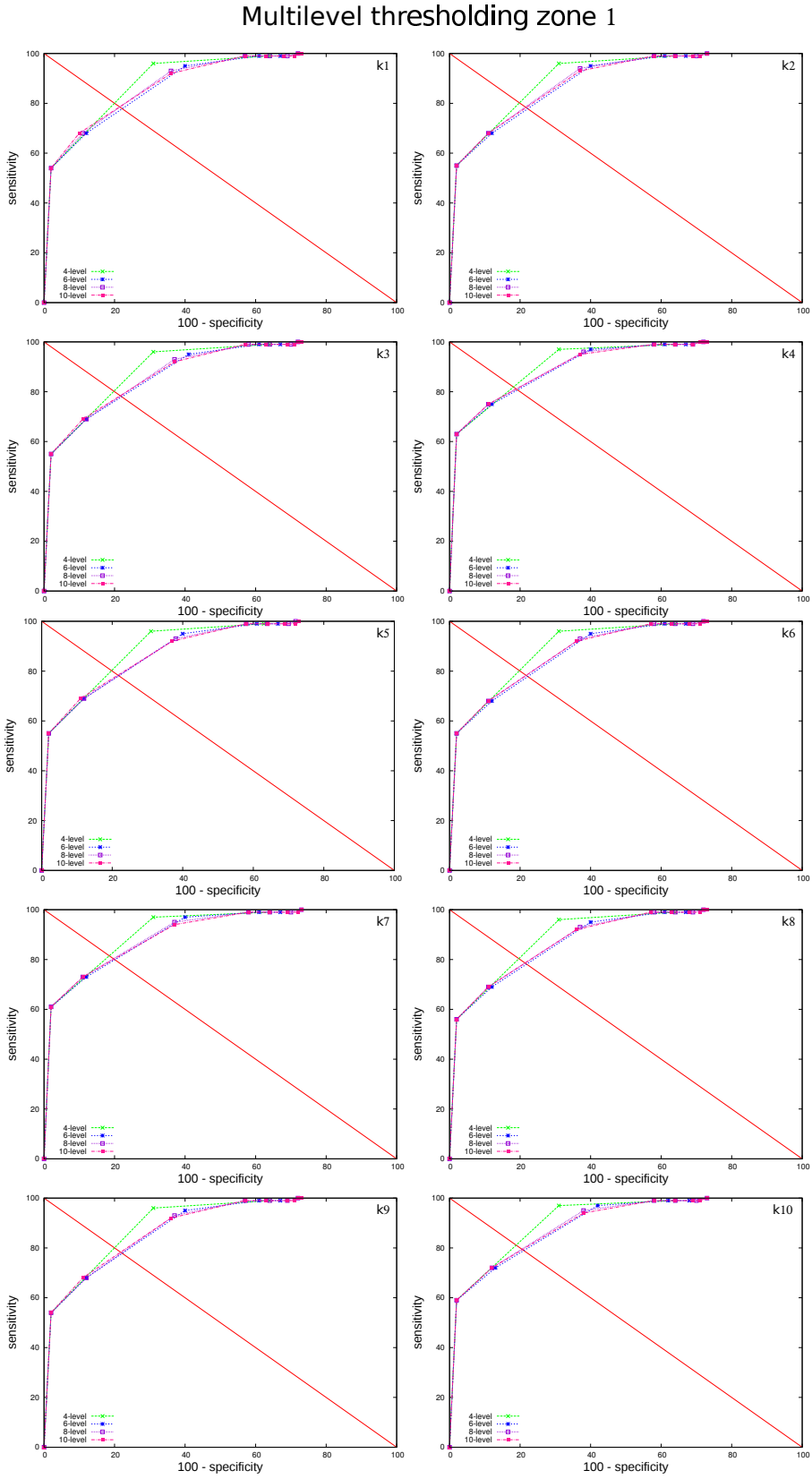


Figure C.5: ROC curves for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding at zone 1 of ERF. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

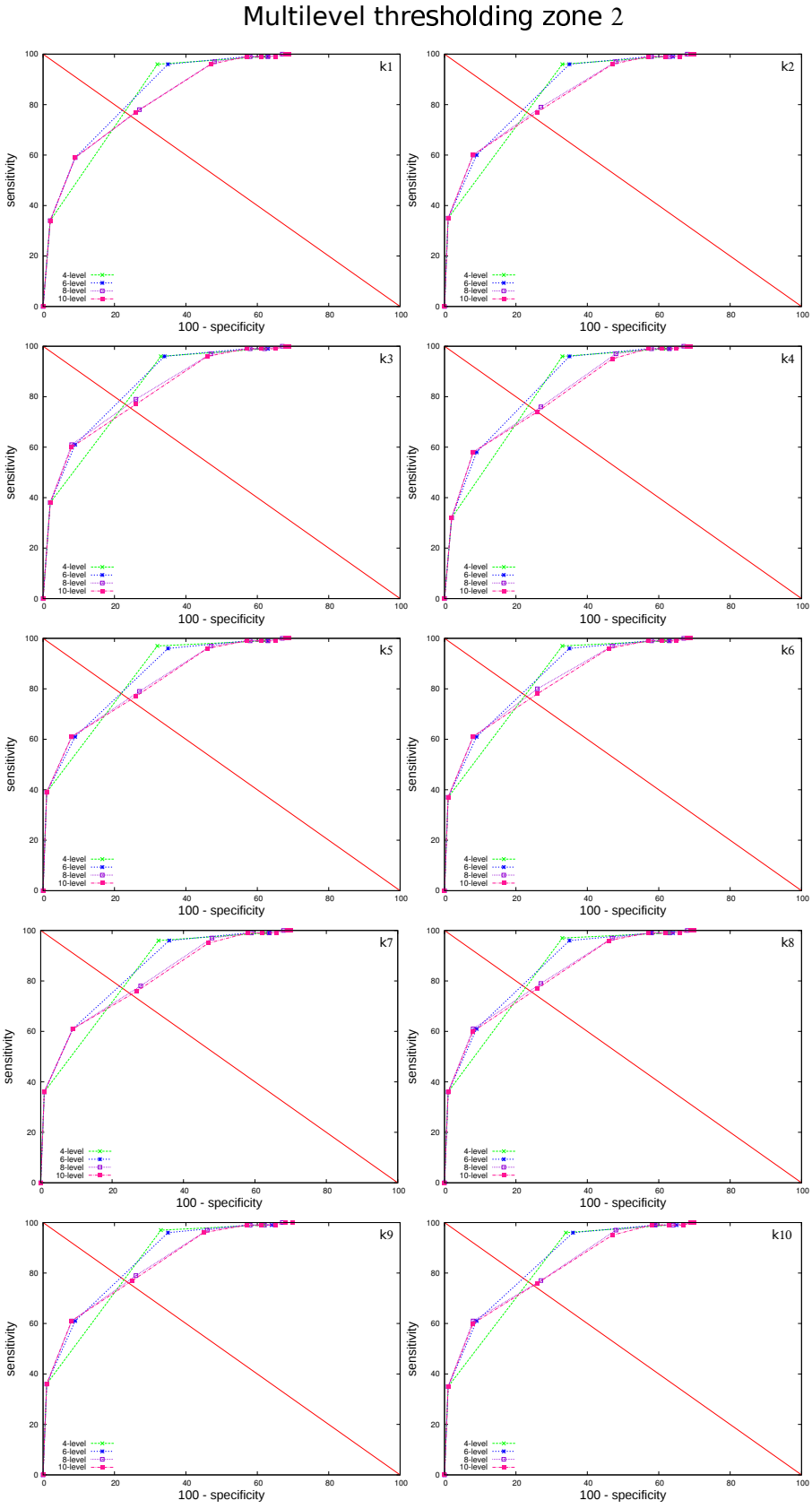


Figure C.6: ROC curves for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding at zone 2 of ERF. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

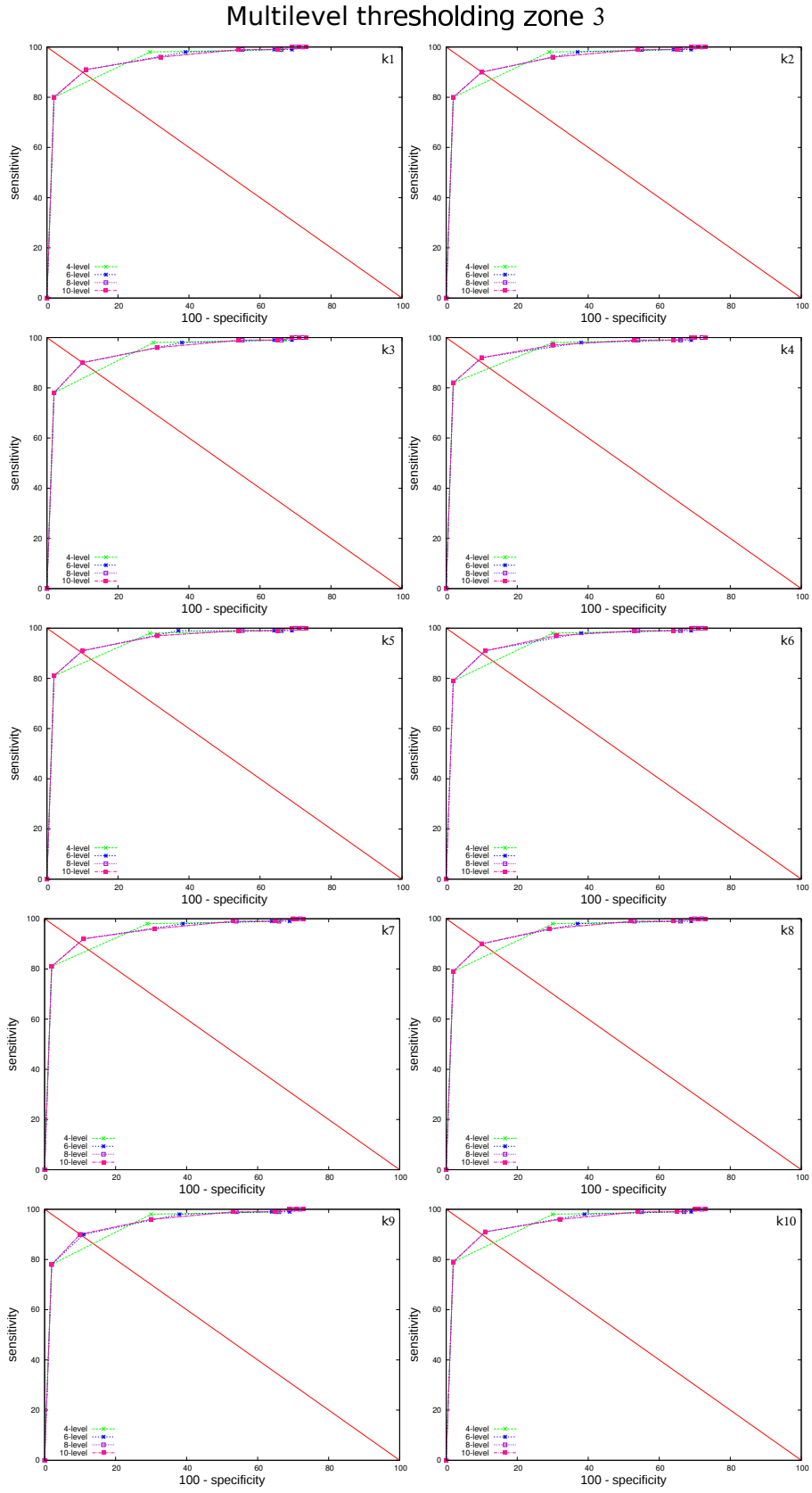


Figure C.7: ROC curves for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding at zone 3 of ERF. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

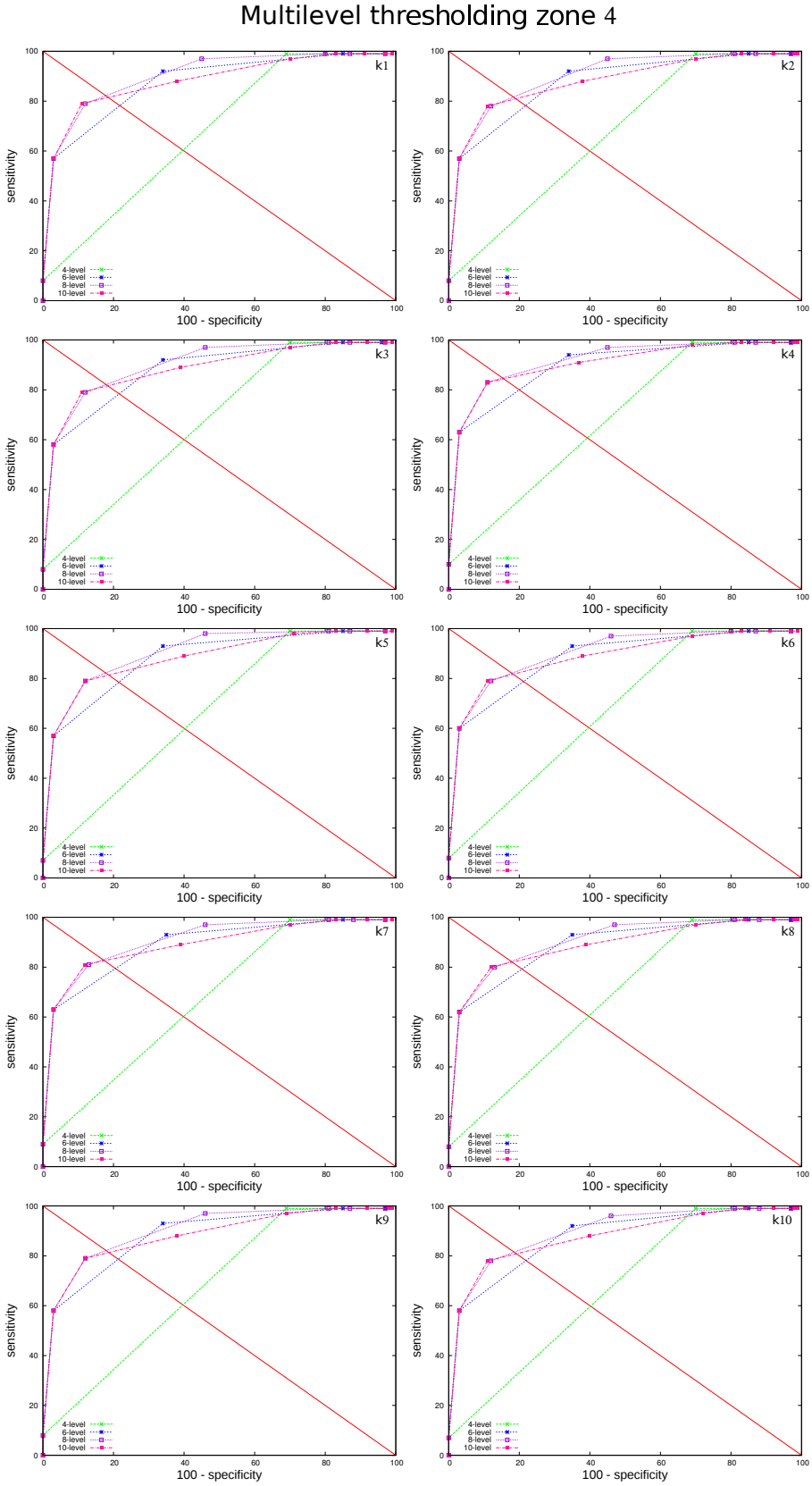


Figure C.8: ROC curves for the 10 k-fold training sets obtained with values of l_b from the levels of $4-level$, $6-level$, $8-level$, and $10-level$ thresholding at zone 4 of ERF. Sensitivity and specificity were computed by comparing the matches between the automatic segmentation and the intersection of the manual annotations performed by both experts.

C.4 Local BUT

The BUT measurement in the local analysis is computed from the evolution curves related to each local zone. To this end, five specific thresholds are computed from a percentage p_e of the maximum height of each local curve.

Two alternatives have been proposed for computing the local BUT measurement. In the first approach, the same p_e is applied to all the local evolution curves. Figure C.9 shows the ROC curves built with p_e taking values from 5% to 95% for the 10 k-fold training sets. The best value for p_e is always 95% of the total height of the evolution curve. The second approach computes a different p_e each local zone, comparing the BUT measurement obtained at that zone with the experts' average. Figures C.10, C.11, C.12, C.13, and C.14 show the ROC curves for the 10 k-fold training sets built with p_e taking values from 5% to 95% of the maximum height of the local evolution curves from zones 0 to 4, respectively.

Once the best values are set, their accuracy is computed in the 10 k-fold test sets. Table C.4 contains the accuracy results of each k-fold test set.

Table C.4: Accuracy (%) of local BUT measurement for the 10 k-fold test sets with the best value for p_e.

	Local ($p_e = 95\%$)	Zone 0 ($p_e = 65\%$)	Zone 1 ($p_e = 90\%$)	Zone 2 ($p_e = 70\%$)	Zone 3 ($p_e = 55\%$)	Zone 4 ($p_e = 75\%$)
k1	91.99	80.50	92.14	94.84	78.13	90.09
k2	94.14	77.42	93.21	89.35	74.65	85.07
k3	84.36	75.86	84.11	80.09	73.61	80.74
k4	88.17	76.79	87.20	83.97	80.81	85.09
k5	93.36	76.08	92.98	84.60	79.09	86.19
k6	92.00	81.11	90.01	91.12	80.02	84.73
k7	88.32	86.74	88.88	84.62	87.55	77.92
k8	90.41	80.73	90.55	87.30	81.41	83.56
k9	86.16	83.76	88.68	80.18	83.89	79.98
k10	92.52	81.57	92.40	88.11	82.29	80.36
Mean	90.14	80.06	90.02	86.42	80.15	83.37

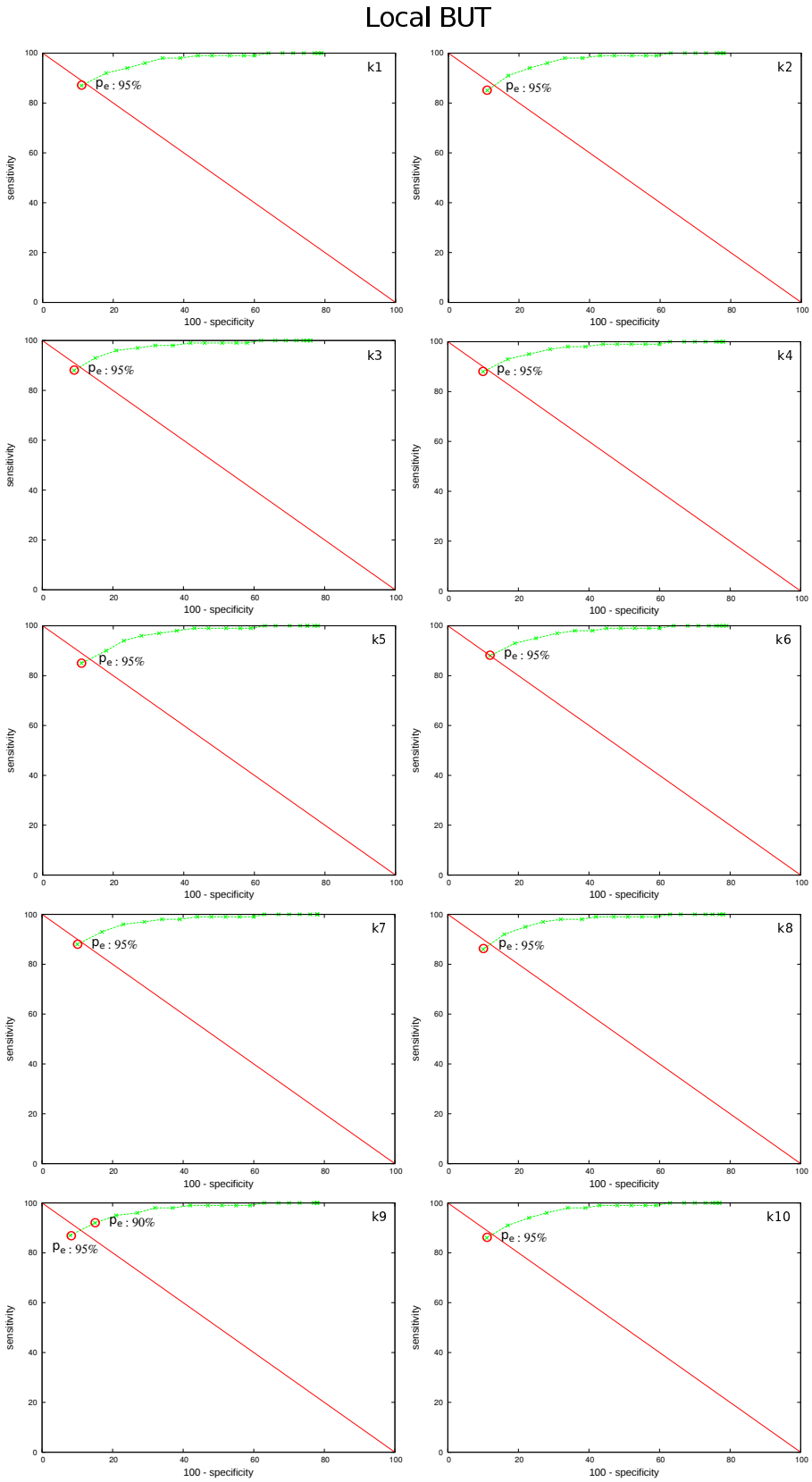


Figure C.9: ROC curves for the 10 k-fold training sets obtained with percentages p_e from 5% to 95% of the maximum height of the local evolution curves. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

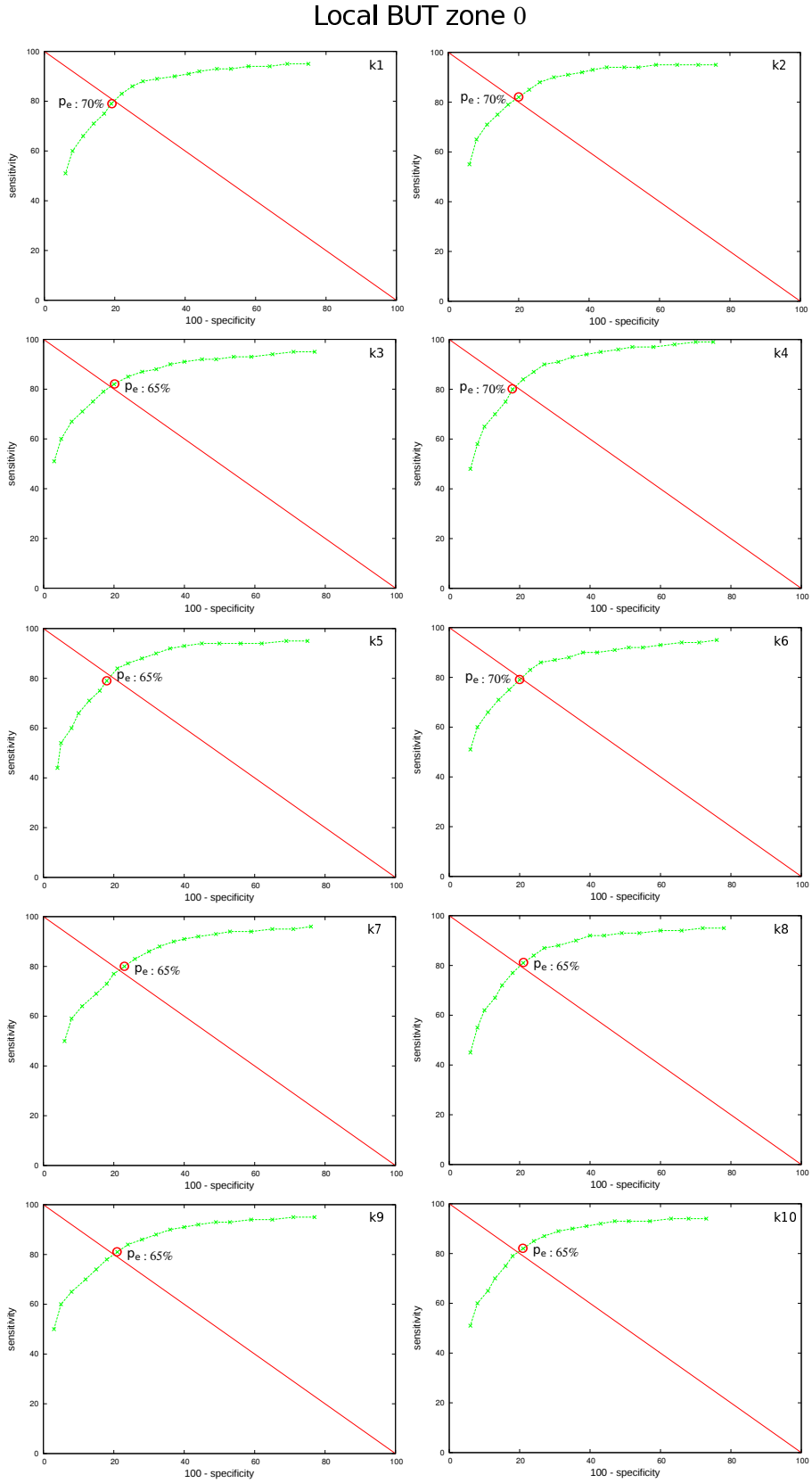


Figure C.10: ROC curves for the 10 k-fold training sets obtained with percentages p_e from 5% to 95% of the maximum height of the evolution curves related to zone 0. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

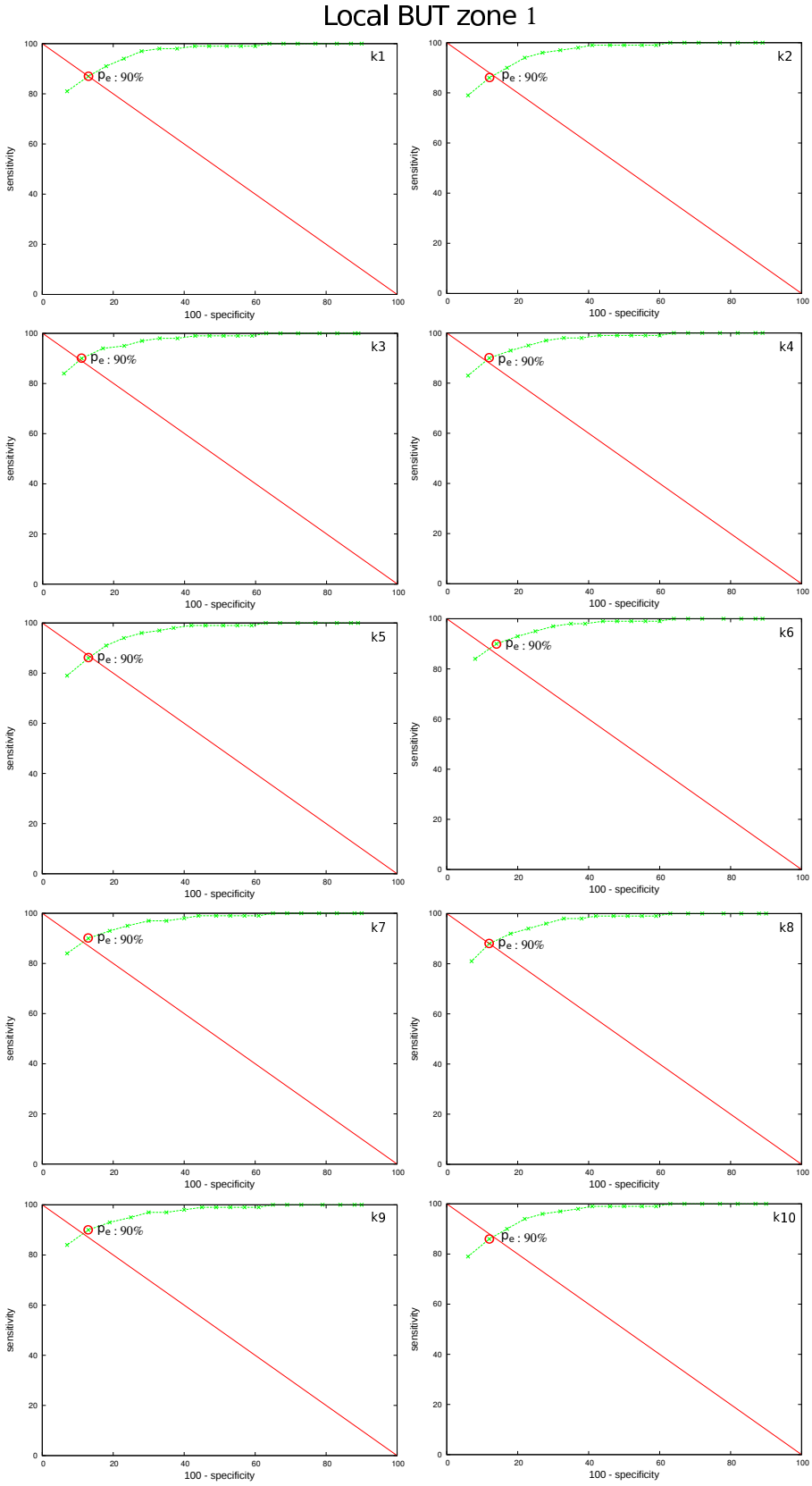


Figure C.11: ROC curves for the 10 k-fold training sets obtained with percentages p_e from 5% to 95% of the maximum height of the evolution curves related to zone 1. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

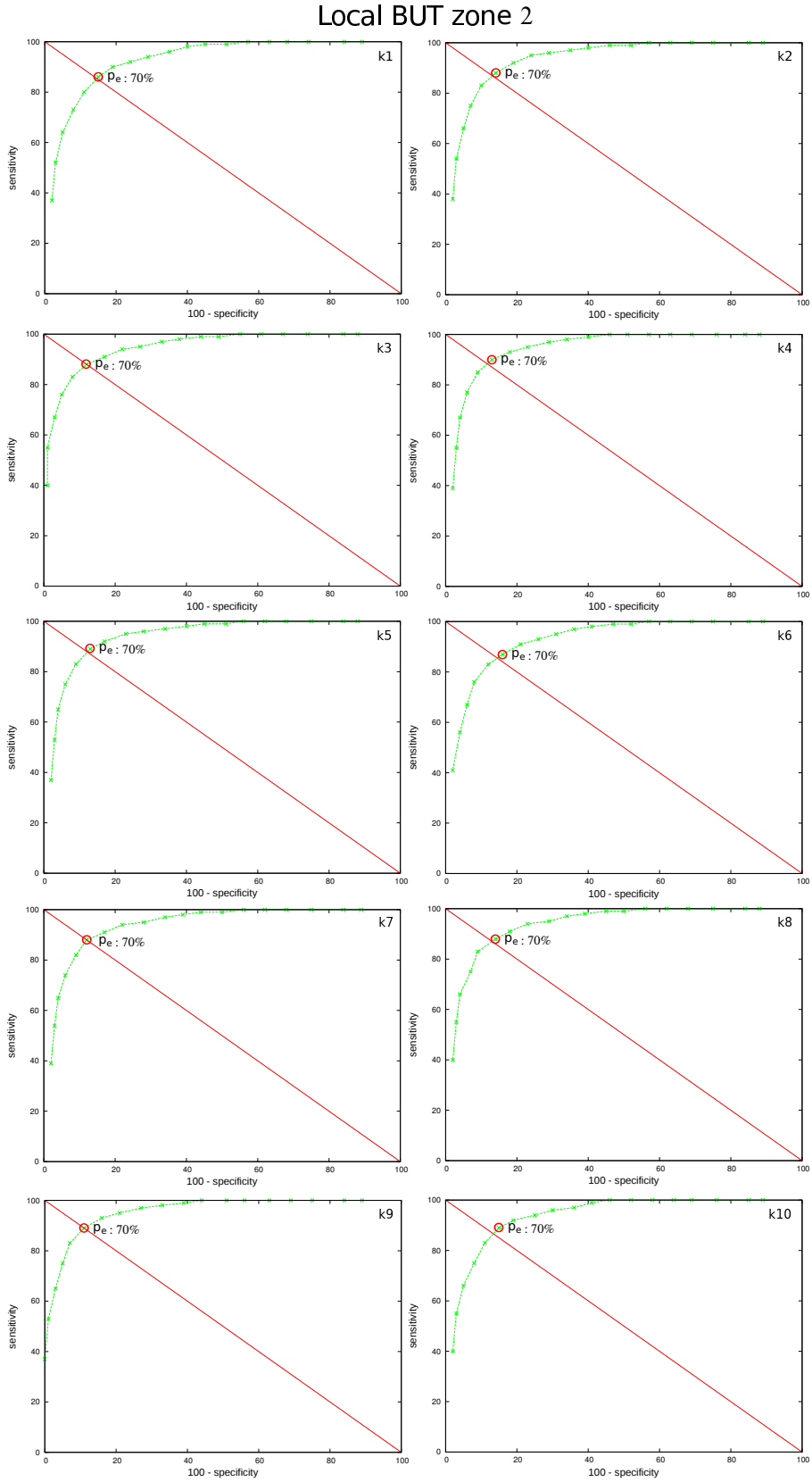


Figure C.12: ROC curves for the 10 k-fold training sets obtained with percentages p_e from 5% to 95% of the maximum height of the evolution curves related to zone 2. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

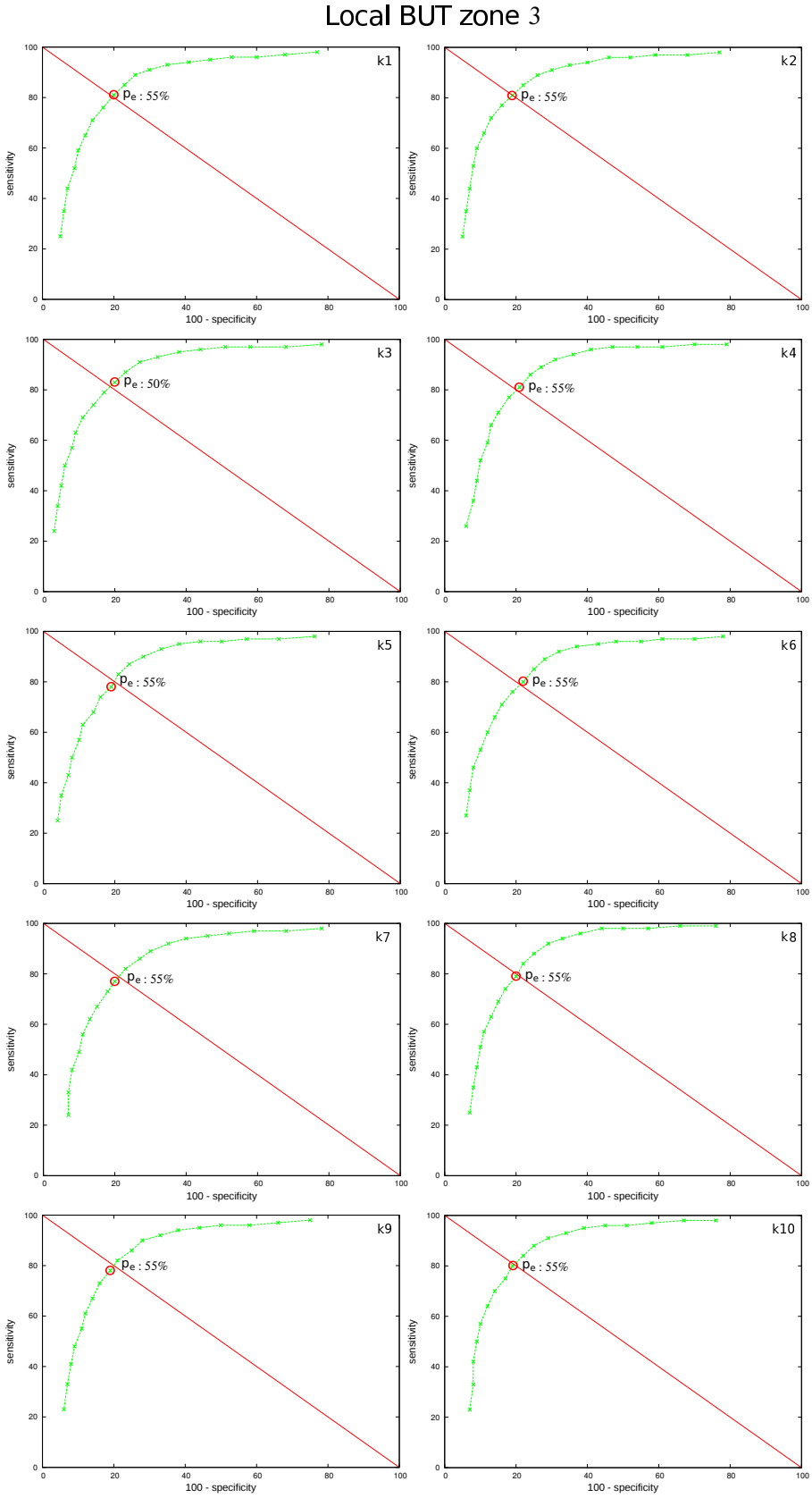


Figure C.13: ROC curves for the 10 k-fold training sets obtained with percentages p_e from 5% to 95% of the maximum height of the evolution curves related to zone 3. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

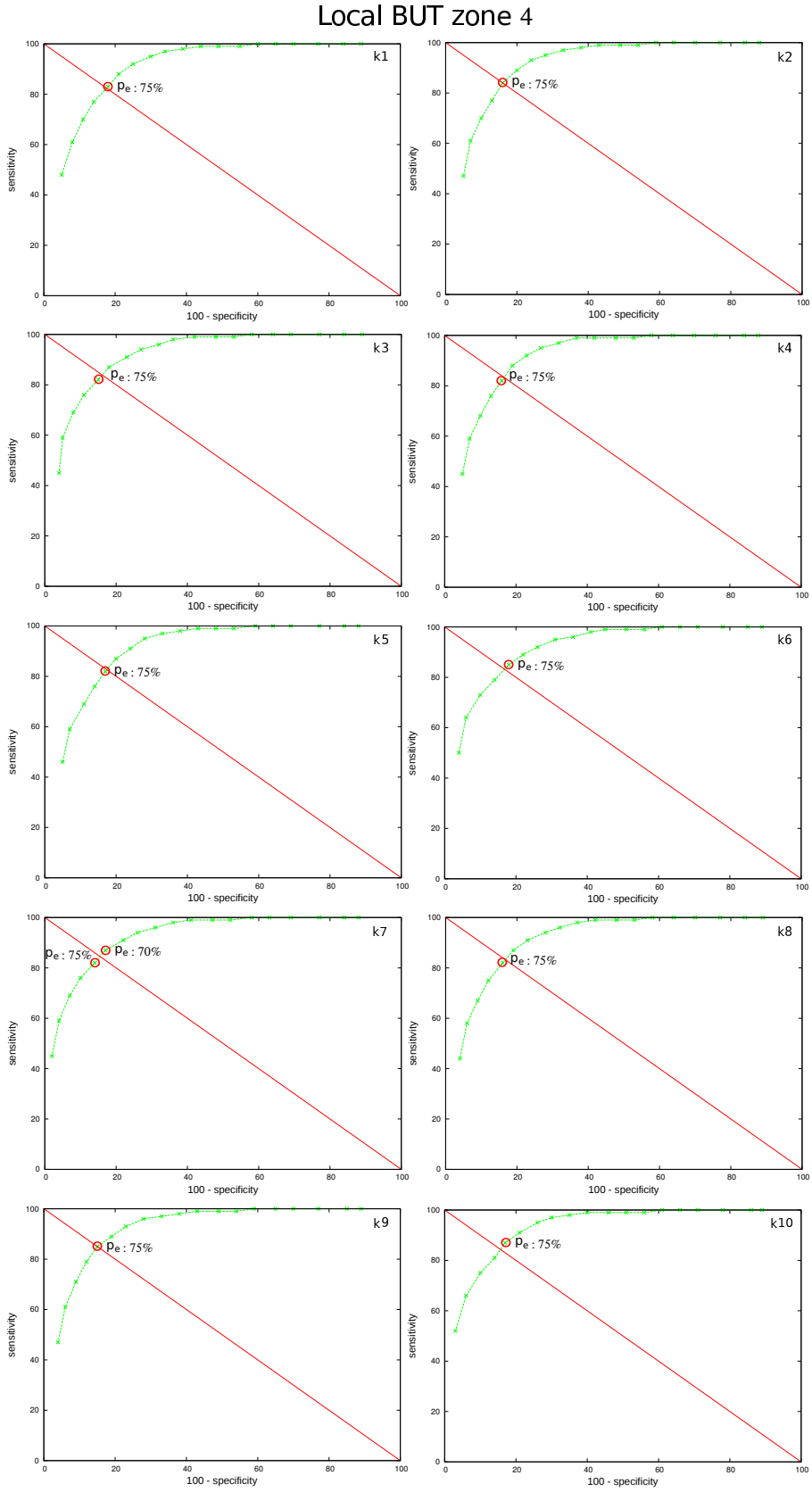


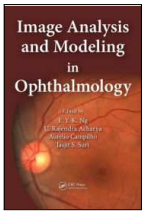
Figure C.14: ROC curves for the 10 k-fold training sets obtained with percentages p_e from 5% to 95% of the maximum height of the evolution curves related to zone 4. Sensitivity and specificity were computed by comparing the matches between the expert average and the BUT measurements computed by the system.

Appendix D

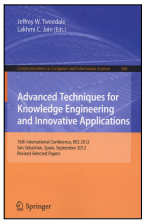
Publications

D.1 Books and Journals

L. Ramos, N. Barreira, A. Mosquera, M. G. Penedo, E. Yebra-Pimentel, C. García-Resúa, "Analysis of parameters for the automatic computation of the tear film Break-Up Time test based on CCLRU standards", Computer Methods and Programs in Biomedicine, 113, 715-724, 2014.



M. G. Penedo, B. Remeseiro, L. Ramos, N. Barreira, C. García-Resúa, E. Yebra-Pimentel, A. Mosquera, "Automatization of Dry Eye Syndrome Tests", Image Analysis and Modeling in Ophthalmology, Chapter 16, 293-320, 2014.



L. Ramos, N. Barreira, A. Mosquera, M. Currás, H. Pena-Verdeal, M. J. Giráldez, M. G. Penedo, "Computational Approach for Measuring the Tear Film Break-Up Time in an unsupervised Manner", Advanced Techniques for Knowledge Engineering and Innovative Applications, 246, 254-267, 2013.

L. Ramos, N. Barreira, A. Mosquera, M. Currás, H. Pena-Verdeal, M.J. Giráldez, M. G. Penedo, "Adaptive parameter computation for the automatic measure of the Tear Break-Up Time", Frontiers in Artificial Intelligence and Applications, 243, 1370 - 1379, 2012.

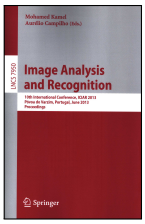
D.2 International Conferences

L. Ramos, N. Barreira, A. Mosquera, H. Pena-Verdeal, M. G. Penedo, "Automatic classification of tear film rupture areas", Viser 2014, November 2014.

L. Ramos, N. Barreira, H. Pena-Verdeal, M. J. Giráldez, "Automatic assessment of tear film break-up dynamics", Studies in Health Technology and Informatics: Innovation in Medicine and Healthcare 2014 (InMed-14), 207, 173-182, San Sebastian, July 2014.

L. Ramos, N. Barreira, J.M. Molinelli, H. Pena-Verdeal, M. J. Giráldez, "Computerized tool for an automatic break-up assessment", II International Conference on Applications of Optics and Photonics (AOP 2014), 43, May 2014.

L. Ramos, N. Barreira, A. Mosquera, H. Pena-Verdeal, M. J. Giráldez, "Computerized tool for evaluating tear film break-up dynamics", International Congress of Optometry and Vision Science (CIOCV 2014), 70, Braga, Portugal, May 2014.



L. Ramos, N. Barreira, A. Mosquera, H. Pena-Verdeal, E. Yebra-Pimentel, "Break-up analysis of the tear film based on time, location, size and shape of the rupture area", LNCS: Image Analysis and Recognition (ICIAR 13), 7950, 695-702, 2013.

L. Ramos, N. Barreira, A. Mosquera, H. Pena-Verdeal, M.J. Giráldez, "Computerized tool for video analysis in a tear film break-up test framework ", VII Iberoamerican Conference on Optics - XI Latinamerican meeting on Optics, Lasers and Applications (RIAO-OPTILAS 2013), Porto, Portugal, July 2013.

L. Ramos, N. Barreira, A. Mosquera, H. Pena-Verdeal, M.J. Giráldez, "Computer Aided Diagnosis system for the evaluation of the tear film Break-Up Test and the characterization of the rupture areas ", International Conference of Optometry and Vision Science (CIOCV 2013), 50, Braga, Portugal , April 2013.

E. Cebreiro, L. Ramos, A. Mosquera, N. Barreira, M. G. Penedo, "Automation of the Tear Film Break-Up Time Test", 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL), Barcelona, October 2011.

D.3 National Conferences

B. Remeseiro, L. Ramos, N. Barreira, A. Mosquera, M. G. Penedo, "Automatización de tests para el diagnóstico del síndrome de ojo seco", BioIntregaSaúde (BIS 2013), 49, Santiago de Compostela, Spain, April 2013.

D.4 Pending of publication

L. Ramos, N. Barreira, H. Pena-Verdeal, M.J. Giráldez, "Computational approach for tear film assessment based on break-up dynamics", Biosystems Engineering , 2014.

D.5 Software Licenses

Herramienta para la medición automática del tiempo de ruptura de la película lagrimal. Autorship shared by Universidade da Coruña, Manuel Francisco González Penedo, Marcos Ortega Hortas, Noelia Barreira Rodríguez and Lucía Ramos García.

Herramienta para la clasificación automática de la ruptura de la película lagrimal. Autorship shared by Universidade da Coruña, Manuel Francisco González Penedo, Marcos Ortega Hortas, Noelia Barreira Rodríguez and Lucía Ramos García.

Appendix E

Resumen

El ojo es uno de los órganos más complejos y sensibles que poseemos. Es la base de la vista y su función consiste en transformar la energía de la luz en señales eléctricas que se envían al cerebro a través del nervio óptico. Como una ventana al mundo, los ojos recogen las imágenes que nos mantienen en contacto con la realidad, siendo responsables de cuatro quintas partes de toda la información que recibe nuestro cerebro.

El globo ocular está en contacto con el medio ambiente a través de la superficie ocular, integrada por la conjuntiva y la cornea. La córnea es una estructura transparente y resistente que ayuda a enfocar la luz entrante, siendo responsable de dos terceras partes de la potencia total del ojo. Hay más terminaciones nerviosas en la córnea que en cualquier otra parte del cuerpo, por lo que es extremadamente sensible. Para permanecer saludable, las células del epitelio corneal deben mantenerse húmedas. Esta necesidad se satisface mediante la película lagrimal, una estructura compleja y dinámica de lípidos, proteínas y mucinas que cubre la superficie anterior de la córnea, protegiéndola y manteniéndola lubricada. La película lagrimal juega un papel esencial para garantizar una visión óptica de buena calidad. Cambios cuantitativos o cualitativos en la película lagrimal suponen un efecto negativo en la calidad de la

visión pudiendo desencadenar en alteraciones como el síndrome de ojo seco.

Es síndrome de ojo seco es un trastorno común de la superficie ocular que afecta a un sector significativo de la población. La prevalencia de esta enfermedad ha ido aumentando en los últimos años, llegando a afectar hasta el 10 - 15% de la población general, y 18 - 30% en los usuarios de lentes de contacto. Diversos factores tales como condiciones ambientales adversas, uso de ciertos medicamentos o tareas visuales que reducen la frecuencia de parpadeo han contribuido a este incremento. Esta enfermedad supone un problema creciente de salud publica que tiene un fuerte impacto en la calidad de vida. Este síndrome limita y degrada la visión presentando síntomas de incomodidad, dolor, irritación, visión borrosa, fatiga ocular y pérdida de la sensibilidad y el contraste. Estos síntomas se correlacionan con la dificultad para trabajar o realizar actividades de la vida diaria tales como la lectura, el uso del ordenador, ver la televisión o conducir.

E.1 Ruptura de la Película Lagrimal

El diagnóstico del síndrome de ojo seco es difícil debido a su etiología multifactorial, por lo que hay varias pruebas clínicas para evaluar diferentes aspectos de la película lagrimal. Una de esas pruebas es el test de BUT, que consiste en medir el tiempo que la película lagrimal permanece estable sin parpadear. Para llevar a cabo esta prueba se instila fluoresceína sódica en el ojo del paciente y se observa la película lagrimal con la ayuda de un filtro azul cobalto acoplado a un biomicroscopio con lámpara de hendidura y un filtro amarillo para mejorar la visibilidad de la emisión de la fluoresceína. Se le pide al paciente que parpadee tres veces de forma natural y sin parpadear, con el fin de distribuir la fluoresceina sobre la cornea, y a continuación, que mantenga el ojo abierto durante el mayor tiempo posible. A medida que transcurre el tiempo la

película lagrimal pierde estabilidad y aparecen manchas oscuras relacionadas con la evaporación de agua sobre la superficie ocular. El test de BUT mide el tiempo transcurrido entre el último parpadeo y la primera aparición de estas manchas que se corresponden con la ruptura de la película lagrimal. Un tiempo de BUT bajo está asociodo a una humectación de la superficie ocular limitada y es uno de los principales signos de una película lagrimal anormal.

El test de BUT solo examina la aparición del primer punto oscuro en la película lagrimal, sin tener en cuenta como evoluciona la zona de ruptura a partir de ese momento. Sin embargo, existen características de la ruptura que podrían estar relacionadas con aspectos específicos de la película lagrimal, pudiendo afectar a los síntomas del síndrome de ojo seco.

Por un lado, la localización de la ruptura puede estar relacionada con la estabilidad de la película lagrimal. En este sentido, la parte central de la córnea es mucho más sensible que la periferia, por lo que esta área es mucho más vulnerable a alteraciones de la película lagrimal. Por tanto, una ruptura situada en el centro podría afectar a la visión como resultado de cambios de la película lagrimal sobre la pupila. Además, la ruptura se produce a veces en la misma posición en diferentes test de BUT del mismo paciente, lo que podría indicar una zona debilitada de la superficie ocular.

Por otra parte, la primera ruptura puede aparece con un pequeño punto o como un área más grande, y puede aumentar a medida que pasa el tiempo, expandiéndose según diferentes patrones de ruptura. En la bibliografía se han identificado tres patrones diferentes denominados como puntos, lineas y regiones. Estos patrones están relacionados con el tiempo de BUT, el tamaño de la ruptura y su tasa de expansión, por lo que son un buen indicador de la estabilidad y calidad de la película lagrimal. Adicionalmente, los patrones de ruptura están asociados con la resistencia a la contaminación mucínica, por lo que también pueden ser un indicador clínico útil para la identificación de

candidatos pobres para el uso de lentes de contacto.

Esta información se omite en el test de BUT original, pero podría ser relevante en la evaluación de la película lagrimal y para la comprensión de su inestabilidad en el síndrome de ojo seco y su relación con los síntomas de la superficie ocular.

E.2 Metodología

El test de BUT está afectado por una baja repetibilidad debido principalmente a la apreciación subjetiva de las zonas oscuras, las diferencias entre los expertos y la variabilidad de la película lagrimal. Además de la subjetividad, la caracterización a mano de la zonas de ruptura es una tarea tediosa que supone mucho tiempo y esfuerzo para los expertos. Por tanto, la automatización de este análisis reduciría su carácter subjetivo permitiendo una evaluación más precisa de la calidad y estabilidad de la película lagrimal, esencial para el diagnóstico del síndrome de ojo seco. El objetivo de este trabajo es desarrollar una metodología automática para la evaluación de la ruptura de la película lagrimal sobre la superficie de la córnea. Esta metodología incluye el cálculo del tiempo de BUT así como un estudio cuantitativo y cualitativo de las zonas de ruptura en base a su localización, tamaño, forma y evolución.

Cada vídeo de la película lagrimal contiene varias secuencias de BUT, denominadas como secuencias de interés (SOIs), que consisten en un conjunto de imágenes delimitadas por parpadeos en los que puede ser llevado a cabo el test de BUT. Así, tras el parpadeo que marca el inicio de la SOI la fluoresceína se distribuye de forma uniforme sobre la película lagrimal. A medida que pasa el tiempo, la película lagrimal va evaporándose y perdiendo estabilidad y aparecen áreas oscuras que se corresponden con su ruptura. Esas zonas de ruptura van evolucionando hasta el parpadeo que marca el final de la SOI. La región de

interés (ROI) para el análisis de la ruptura se corresponde con la parte visible del iris, que puede variar ligeramente a lo largo de la secuencia, dependiendo del grado de apertura del ojo y de la presencia de sombras debidas a partes externas del ojo como los párpados y las pestañas.

Así, en el Capítulo 2 se describe la primera etapa de la metodología que consiste en un preprocesado de los vídeos de película lagrimal para localizar las diferentes SOIs y extraer la ROI en cada imagen a lo largo de las secuencias. Para delimitar las SOIs se detectan los parpadeos a partir de las diferencias finitas de los valores medios de gris entre imágenes consecutivas. A continuación, la ROI de cada imagen es extraída mediante correlación con un conjunto de máscaras elípticas en el dominio de la frecuencia. Después las ROIs son alineadas a lo largo de la SOI para que la metodología sea robusta ante ligeros movimientos del ojo. Además, se lleva a cabo un ajuste adaptativo de la ROI según las características del ojo. Así, los radios de la elipse que representa el iris se reducen ligeramente para eliminar zonas difusas en el borde y la ROI se corta por la parte superior e inferior para descartar partes externas del ojo como párpados y pestañas.

Una vez que se han delimitado las diferentes SOIs y la ROI ha sido extraída en cada imagen, la etapa principal de la metodología es la evaluación de la ruptura de la película lagrimal, que consiste en analizar cada SOI para determinar el tiempo de BUT así como caracterizar las diferentes zonas de ruptura. En el Capítulo 3 se detallan los diferentes pasos para este análisis. En primer lugar se lleva a cabo un preprocesado para normalizar la variabilidad de la luz y el contraste en las imágenes de la película lagrimal. Después se calcula un umbral de ruptura en cada SOI, es decir, la máxima intensidad para considerar que un píxel forma parte de la zona de ruptura.

Después del paso de preprocesado se lleva a cabo el análisis de la ruptura en cada SOI. Para esto, en primer lugar se calcula la medida de BUT como

el tiempo transcurrido desde el principio de la SOI hasta que se detecta la aparición de zonas oscuras, es decir, cuando ocurre la ruptura de la película lagrimal. Las zonas de ruptura pueden evolucionar con el tiempo, por lo que son analizadas desde su primera aparición hasta el final de la SOI. Por un lado se obtiene un descriptor para cada zona de ruptura en base a diferentes características morfológicas y se aplica un clasificador para decidir el patrón de ruptura al que pertenece cada una de las zonas segmentadas. Además, se mide el tamaño y tasa de crecimiento de las zonas de ruptura a lo largo de la SOI que permiten caracterizar la dinámica de la ruptura y extraer información adicional.

Por otro lado, la ubicación de la ruptura también está relacionada con aspectos específicos de la película lagrimal y la severidad del ojo seco. Así, la metodología descrita es extendida para llevar a cabo un análisis local de la ruptura que permita una caracterización más específica. Con este objetivo, la ROI se divide en 5 partes iguales según el estándar CCLRU (Cornea and Contact Lens Research Unit), y cada una de las características consideradas en el análisis global es evaluada por separado en cada una de las zonas locales.

En el Capítulo 4 se resumen los experimentos llevados a cabo para validar la metodología propuesta y los resultados extraídos.

Finalmente el Capítulo 5 resume las principales aportaciones del trabajo así como futuras líneas de investigación.

E.3 Conclusiones

En este trabajo se ha desarrollado una metodología nueva para una evaluación de la ruptura de la película lagrimal totalmente automática. Esta metodología permite un análisis cuantitativo y cualitativo de la estabilidad de la película lagrimal, como extensión al test de BUT.

De acuerdo con la validación realizada en el Capítulo 4 se pueden hacer las siguientes observaciones. Por un lado, la localización de las diferentes SOIs y la extracción de la ROI en cada imagen de las SOIs presentan un buen rendimiento, obteniendo resultados de precisión por encima del 90% que permiten definir correctamente el espacio de análisis de la ruptura de la película lagrimal.

Con respecto a la segmentación de la ruptura hay que tener en cuenta que el marcado manual de las zonas de ruptura es una tarea tediosa y subjetiva, existiendo un desacuerdo considerable entre los expertos. La coincidencia de las zonas de ruptura delimitadas a mano por ambos expertos está en torno al 60%, por lo que se usa la intersección entre ambos expertos para la validación. La metodología propuesta proporciona resultados del 90% de precisión, permitiendo extraer una buena representación de las zonas de ruptura de película lagrimal.

En cuanto a la medida del tiempo de BUT, la variabilidad entre los expertos está en un intervalo de ± 2.5 segundos, debido también a la subjetividad de la apreciación de la ruptura. El cálculo global de BUT obtiene resultados de precisión en torno al 88% con una desviación media de 1.21 segundos. Este valor concuerda con la medida manual ya que está dentro del mismo rango que hay entre los propios expertos.

La clasificación de las zonas de ruptura presenta una precisión global del 85% para todos los patrones. En algunos casos, zonas pequeñas que presentaban patrón región fueron clasificadas con patrón punto, así como algunos puntos bastante grandes fueron clasificados como regiones. Algunas zonas con patrón línea no muy definida fueron confundidas con patrones región mientras que otras zonas pequeñas con patrón línea fueron clasificadas con patrón punto. Además, la caracterización de la dinámica de la ruptura permite distinguir entre una ruptura temprana o tardía, si la primera ruptura representa un punto

pequeño o una zona amplia de la ROI y si se expande lenta o rápidamente.

Finalmente, el análisis local ha permitido una caracterización más específica centrada en las zonas de ruptura. La segmentación llevada a cabo en las diferentes zonas locales ha proporcionado resultados similares a los obtenidos en la segmentación global. Sin embargo, la medida de BUT ha obtenido resultados ligeramente mejores, con un porcentaje del 95.55% dentro del rango de ± 2.5 segundos existente entre los propios expertos y con una desviación media de 0.96 segundos.

Los experimentos llevados a cabo en este trabajo han demostrado que esta metodología proporciona resultados que se encuentran en los mismos rangos que los propios expertos. Estos demuestran que la automatización del análisis de la ruptura lagrimal no sólo es factible sino que también ofrece nuevas oportunidades para el diagnóstico y seguimiento del síndrome de ojo seco. El sistema automático proporciona resultados rápidos y objetivos, así como supone un ahorro a los expertos en esfuerzo, tiempo y costes. Además, este sistema ofrece la posibilidad de llevar a cabo grandes estudios poblacionales para extraer nuevas variables de interés para la práctica clínica.

Bibliography

M. Abelson, G. Ousler 3rd, L. Nally, D. Welch, and K. Krenzer. Alternative reference values for tear film break up time in normal and dry eye populations. *Adv Exp Med Biol*, 506(Pt B):1121–5, 2002.

E. Acharya, U.R. andNg and J. Suri. *Image Modeling of the Human Eye*. Bioinformatics and biomedical imaging. Artech House, 2008.

S. Arora, J. Acharya, A. Verma, and P. K. Panigrahi. Multilevel thresholding for image segmentation through a fast statistical recursive algorithm. *Pattern Recogn. Lett.*, 29(2):119–125, 2008.

C. Begley, N. Himebaugh, H. Renner, D. amd Liu, R. Chalmers, T. Simpson, and J. Varikooty. Tear breakup dynamics: a technique for quantifying tear film instability. *Optom Vis Sci.*, 83:15–21, 2006.

E. Bitton and J. Lovasik. Longitudinal analysis of precorneal tear film rupture patterns. *Adv Exp Med Biol*, 438:381–9, 1998.

A. Bron, F. Sci, and J. Tiffany. The contribution of meibomian disease to dry eye. *Ocul Surf*, 2(2):149–65, 2004.

C. J. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.

J. Canny. A computational approach to edge detection. *IEEE PAMI*, pages 679–698, 1986.

N. Carlson. *Psychology: The Science of Behaviour*. Pearson Education Canada, 2005. ISBN 9780205426867.

R. L. Chalmers, C. G. Begley, and B. Caffery. Validation of the 5Item Dry Eye Questionnaire (DEQ-5): Discrimination across self-assessed severity and aqueous tear deficient dry eye diagnoses. *Contact Lens and Anterior Eye*, 33:55–60, 2010.

T. Chang-Ling. Sensitivity and neural organization of the cat cornea. *Invest Opthalmol Vis Sci.*, 30:1075–82, 1989.

P. Cho. Reliability of a portable noninvasive tear break-up time test on hong kong-chinese. *Optom Vis Sci*, 70:1049–54, 1993.

P. Cho and B. Brown. Review of the tear break-up time and a closer look at the tear break-up time of hong kong chinese. *Optom Vis Sci*, 70:30–8, 1993.

P. Cho et al. Reliability of the tear break-up time thechnique of assessing tear stability and the locations of thear break-up in hong kong chinese. *Optom Vis Sci*, 69:879–85, 1992.

B. Day and P. Brown. Evidence for subcortical involvement in the visual control of human reaching. *Brain*, 124:1832–40, 2001.

B. E. Dougherty, J. J. Nichols, and K. K. Nichols. Rasch analysis of the ocular surface disease index (osdi). *Invest Ophthalmol Vis Sci*, 52(12):8630–5, 2011.

E. Dougherty. *An introduction to morphological image processing*. Tutorial texts in optical engineering. SPIE Optical Engineering Press, 1992.

R. O. Dror, E. H. Adelson, and A. S. Willsky. Surface reflectance estimation and natural illumination statistics. In *IN PROC. OF IEEE WORKSHOP*

ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION, 2001.

M. Elliott, H. Fandrich, and T. Simpson. Analysis of the repeatability of tear break-up time measurement techniques on asymptomatic subjects before, during and after contact lens wear. *Cont Lens Anterior Eye*, 21:98–103, 1998.

T. Fawcett. An introduction to roc analysis. *Pattern Recogn. Lett.*, 27(8): 861–874, June 2006.

C. Fenga, P. Aragona, A. Cacciola, R. Spinela, C. D. Nola, F. Ferreri, and L. Rania. Melbonian gland dysfunction and ocular discomform in video display terminal workers. *Eye*, 22:91–95, 2008.

M. Foracchia, E. Grisan, and A. Ruggeri. Luminosity and contrast normalization in retinal images. *Medical Image Analysis*, 9(3):179–190, 2005.

C. García-Resúa, M. Lira, and E. Yebra-Pimentel. Evaluación superficial en jóvenes universitarios. *Rev. Esp. Contact*, 12:37–41, 2005.

C. García-Resúa, J. Santodomingo-Rubido, M. Lira, M. Giráldez, and E. Yebra-Pimentel. Clinical assessment of the lower tear meniscus height. *Ophthalmic Physiol Opt*, 29:487–496, 2009.

V. K. Gothwal, K. Pesudovs, T. A. Wright, and C. W. McMonnies. Mcmonnies questionnaire: enhancing screening for dry eye syndromes with rasch analysis. *Invest Ophthalmol Vis Sci*, 51(3):1401–7, 2010.

J. Graig and A. Tomlinson. Importance of the lipid layer in human tear film stability and evaporation. *Optom Vis Sci.*, 74:8–13, 2007.

J. Guillon. Non-invasive tearscope plus routine for contact lens fitting. *Cont Lens Anterior Eye*, 21 Suppl 1, 1998.

R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.

M.-K. Hu. Visual pattern recognition by moment invariants, computer methods in image analysis. *IRE Transactions on Information Theory*, 8:179–187, 1962.

M. S. Hughes. Dictionary of eye terminology. *Archives of Ophthalmology*, 109 (9):1208, 1991.

M.-A. Javadi and S. Feizi. Dry eye syndrome. *Journal of Ophthalmic & Vision Research*, 6(3), 2011.

F. V. Jensen. *Introduction to Bayesian Networks*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1st edition, 1996.

M. Johnson and P. Murphy. The effect of instilled fluorescein solution volume on the values and repeatability of tbut measurements. *Cornea*, 24:811–7, 2005.

I. Jolliffe. *Principal Component Analysis*. Springer Verlag, 1986.

S. Kastelan, A. Lukenda, J. Salopek-Rabati, J. Pavan, and M. Gotovac. Dry eye symptoms and signs in long-term contact lens wearers. *Coll Antropol*, 37 Suppl 1:199–203, 2013.

S. Khanal. Dry eye diagnosis. *Invest Opthalmol Vis Sci.*, 49:1907–14, 2008.

P. King-Smith, B. Fink, N. JJ., R. Braun, and G. McFadden. The contribution of lipid layer movement to tear film thinning and breakup. *Invest Opthalmol Vis Sci.*, 50:2747–56, 2009.

D. Korb. *The Tear Film: Structure, Function and Clinical Examination*. Butterworth-Heinemann, 2002. ISBN 9780750641968.

D. Korb, J. Greiner, and J. Herman. Comparison of fluorescein break-up time measurement reproducibility using standard fluorescein strips versus the dry eye test (det) method. *Cornea*, 20:811–5, 2001.

J. Larke. *The Eye in Contact Lens Wear*. Butterworth-Heinemann, 1997. ISBN 9780750644389.

M. J. H. Lee and M. C. W. Kee. The significance of tear film break-up time in the diagnosis of dry eye syndrome. *Kor J. Opthalmol*, 21:69–71, 1988.

M. Lemp. Advances in understanding and managing dry eye disease. *Am. J. Ophthalmol.*, 46:350–356, 2008.

M. Lemp. Report of the national eye institute/industry workshop on clinical trials in dry eyes. *CLAO J*, 21(4):221–32, 1995.

M. Lemp, C. Baudouin, J. Baum, M. Dogru, G. Foulks, S. Kinoshita, P. Laibson, J. McCulley, J. Murube, S. Pfugfelder, M. Rolando, and I. Toda. The definition and classification of dry eye disease: Report of the definition and classification subcommittee of the international dry eye workshop (2007). *Ocul Surf*, 5(2):75–92, 2007.

H. Liu, C. Begley, R. Chalmers, G. Wilson, S. Srinivas, and J. Wilkinson. Temporal progression and spatial repeatability of tear breakup. *Optom Vis Sci.*, 83:723–30, 2006.

G. Lowther. Dryness, tears and contact lens wear. *Usa: reed elsevier*, 1977.

S. McGinnigle, S. Naroo, and F. Eperjesi. Evaluation of dry eye. *Survey of ophthalmology*, 57(4):293–316, 2012.

C. W. McMonnies and A. Ho. Patient history in screening for dry eye conditions. *J Am Optom Assoc*, 58(4):296–301, 1987a.

C. W. McMonnies and A. Ho. Responses to a dry eye questionnaire from a normal population. *J Am Optom Assoc*, 58(7):588–91, 1987b.

B. Miljanovic, R. Dana, D. A. Sullivan, and D. A. Schaumberg. Impact of dry eye syndrome on vision-related quality of life. *Am J Ophthalmol*, 143(3): 409–15, 2007.

R. Montes-Mico and W. Alio, J.L. andCharman. Dynamic changes in the tear film in dry eyes. *Invest Opthalmol Vis Sci.*, 46:1615–9, 2005.

K. Nichols, C. Begley, B. Caffery, and L. Jones. Symptoms of ocular irritation in patients diagnosed with dry eye. *Optom Vis Sci*, 76:838–44, 1999.

K. Nichols, J. Nichols, and G. Mitchell. The lack of association between signs and symptoms in patients with dry eye disease. *Cornea*, 23(8):762–70, 2004.

J. F. Nunes, P. Moreira, and J. Tavares. Shape based image retrieval and classification. In *5th Iberian Conference on Information Systems and Technologies (CISTI)*, 2010.

S. Patel, D. Murray, and A. McKenzie. Effects of fluorescein on tear breakup time and on tear thinning time. *Optom Physiol Opt*, 62:188–90, 1985.

D. M. W. Powers. Evaluation: from precision, recall and F-factor to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2:37–63, 2007.

C. Poynton. Frequently asked questions about color, December 1999. URL http://www.poynton.com/notes/colour_and_gamma/ColorFAQ.html. Accessed: 2014-09-30.

P. Reddy, O. Grad, and K. Rajagopalan. The economic burden of dry eye: a conceptual framework and preliminary assessment. *Cornea*, 23(8):751–61, 2004.

M. Reed Teague. Image analysis via the general theory of moments. *Journal of the Optical Society of America (JOSA)*, 70(8):920–930, 1980.

B. Remeseiro, M. Penas, N. Barreira, A. M. González, J. Novo, and C. García-Resúa. Automatic classification of the interferential tear film lipid layer using colour texture analysis. *Computer Methods and Programs in Biomedicine*, 111(1):93–103, 2013.

J. Rodriguez, A. Perez, and J. Lozano. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(3):569–575, 2010.

L. Rokach and O. Maimon. Top-down induction of decision trees classifiers - a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, (4):476–487, 2006.

F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408, 1958.

T. L. Simpson, P. Situ, L. W. Jones, and D. Fonn. Dry eye symptoms assessed by four questionnaires. *Optom Vis Sci*, 85(8):692–9, 2008.

W. Stevenson, S. Chauhan, and R. Dana. Dry eye disease: an immune-mediated ocular surface disorder. *Arch Ophthalmol*, 130(1):90–100, 2012.

S. Suzuki and K. Abe. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 30(1):32–46, 1985.

D. Sweeney, T. Millar, and S. Raju. Current patterns in the use of diagnostic tests in dry eye evaluation. *Cornea*, 27:656–62, 2008.

R. L. Terry, C. M. Schnider, B. A. Holden, R. Cornish, T. Grant, D. Sweeney, D. La Hood, and A. Back. Cclru standards for success of daily and extended wear contact lenses. *Optom Vis Sci*, 70(3):234–43, 1993a.

T. Terry, C. Schnider, and B. Holden. The CCLRU standards. *Optician*, pages 18–23, 1993b.

A. Tomlinson, S. Khanal, K. Ramaesh, C. Diaper, and A. McFadyen. Tear film osmolarity: determination of a referent for dry eye diagnosis. *Invest Ophthalmol Vis Sci*, 47(10):4309–15, 2006.

R. Tutt, A. Bradley, C. Begley, and L. Thibos. Optical and visual impact of tear break-up in human eyes. *Investigative Ophthalmology & Visual Science*, 41:4117–4123, 2000.

S. Vashisht and S. Singh. Evaluation of Phenol Red Thread test versus Schirmer test in dry eyes: A comparative study. *International Journal of Applied and Basic Medical Research*, 1(1):40–42, 2011.

T. Yedidya, R. Hartley, and J. Guillon. Automatic detection of pre-ocular tear film break-up sequence in dry eyes. *DICTA*, pages 442–448, 2008.

T. Yedidya, P. Carr, R. Hartley, and J. Guillon. Enforcing monotonic temporal evolution in dry eye images. *MICCAI*, pages 976–9, 2009.