# Enhanced Assessment of the Wound-Healing Process by Accurate Multiview Tissue Classification

Hazem Wannous, Yves Lucas, Sylvie Treuillet

# Enhanced assessment of the wound-healing process by accurate multi-view tissue classification

Hazem Wannous, Yves Lucas, *Member, IEEE*, and Sylvie Treuillet

*Abstract*—With the widespread use of digital cameras, free-hand wound imaging has become common practice in clinical settings. There is however still a demand for a practical tool for accurate wound healing assessment, combining dimensional measurements and tissue classification in a single user-friendly system. We achieved the first part of this objective by computing a 3D model for wound measurements using uncalibrated vision techniques. We focus here on tissue classification from color and texture region descriptors computed after unsupervised segmentation. Due to perspective distortions, uncontrolled lighting conditions and view points, wound assessments vary significantly between patient examinations. The main contribution of this paper is to overcome this drawback with a multi-view strategy for tissue classification, relying on a 3D model onto which tissue labels are mapped and classification results merged. The experimental classification tests demonstrate that enhanced repeatability and robustness are obtained and that metric assessment is achieved through real area and volume measurements and wound outline extraction. This innovative tool is intended for use not only in therapeutic follow-up in hospitals but also for telemedicine purposes and clinical research, where repeatability and accuracy of wound assessment are critical.

*Index Terms*—multi-view classification, 3D modeling, wound assessment.

1

## I. WOUND ASSESSMENT PRACTISE

### A. Clinical practise

Monitoring the wound healing process is a tedious task for clinicians and nurses as it is necessary to periodically assess the wound. All types of wounds are concerned: not only chronic wounds but also ulcers, burns, traumatic or surgical wounds and dermatological lesions. Moreover, wound care is expensive: according to a report published by the NIGMS in 2008 (*http://www.nigms.nih.gov*) in the USA, chronic wounds cost the nation $20 billion to $25 billion and acute or traumatic wounds add another $7 to $10 billion to the bill annually, as the healing process can last several months;with the ageing of

[1]Hazem Wannous is with the ENSEIRB engineering school, IMS Laboratory, 351 cours de la liberation, 33405 Talence cedex, France (e-mail: hazem.wannous@ims-bordeaux.fr

Yves Lucas is with the University of Orleans, IUT of Bourges, Physical Measurement Dept, PRISME Institute, Image and signal for the system Group, 63 av. de Lattre 18020 Bourges cedex, FRANCE (corresponding author phone +33 248 238 058, e-mail: yves.lucas@bourges.univ-orleans.fr)

Sylvie Treuillet is with the Dept of Signal and Image Processing, Ecole Polytechnique , PRISME Institute, Image and signal for the system Group, University of Orleans, 12 rue de Blois, 45067 Orleans cedex, FRANCE (e-mail: sylvie.treuillet@univ-orleans.fr)

the population this cost will necessarily increase by 25% over the next 10 years.

As health care costs need to be drastically reduced, there is a growing demand for patients to be cared for at home;wound monitoring could be carried out from a distance, outside a hospital environment, in private homes properly equipped for telemedicine practise. Pioneer experiments in this area consisted simply in uploading images to a web site where a physician could view the data at his convenience [30]. In more recent studies, image processing has been added but it provides only ulcer stage grading [26]. The quantitative assessment of chronic wounds still relies on visual inspection and manual techniques to describe the shape of the wound (perimeter, surface, depth ...) and the biological nature of the skin tissues (percentage of each class, wound severity stage, burn degree ...) [46], [42]. Wound dimensions and shape are currently measured with an ordinary ruler, or sometimes through sketches on cross-ruled sheets, serum injection or alginate moldings. Assessing the type and proportion of tissues likewise remains highly empirical as evaluation is performed visually and then recorded on a red-yellow-black scale corresponding respectively to the dominant color of the different tissues found on a wound: granulation, slough and necrosis. Healing is a complex cascade of cellular events operating to reconstruct damaged tissues, and also an individual process that exhibits considerable inter-patient variability. As the different tissues may overlap and be difficult to distinguish, wound assessment is not straightforward. The lack of quantitative data affects the coordination of care staff and hinders clinical studies focused on healing. Digital cameras, though now widespread in clinical centers, are used only for basic patient data recording and not image processing, as wound therapeutic follow-up is mainly carried out by nurses.

### B. Wound imaging studies

While several studies have tackled the problem of wound assessment, these attempts have failed to provide a robust tool for automatic tissue classification. Results remain too dependent on image capture conditions, sample database building, region descriptor selection, tissue class learning protocol etc..., preventing reproducible results from being obtained within the complete image processing chain. Several features contribute to making automatic classification difficult. First, wound image acquisition requires technical skill, especially in the patient room where lighting is not controlled. At close

range, the depth of field remains always limited in macro mode, ambient light is insufficient and may easily result in fuzzy images. Moreover, the patient is rarely able to maintain a convenient posture for a snapshot. After image capture and image compression, an important requirement for data management [44], the wound area must be extracted from the image. Early attempts to accomplish this by using color measurements did not completely succeed, as reported by [23] and have been replaced by semi-automatic methods [35] by spline fitting from a set of clicked boundary points or snakes on a manually specified closed contour [39]. As all the photos are taken with the same camera and lighting during experimentation, color correction is often neglected, making it impossible to process images from other care centers correctly. Sometimes, a white patch is included in the field of view to estimate the white balance [40]. More rarely a color pattern is used for enhanced color correction, giving also access to the scale factor for dimensional measurements [19], [35]. The color analysis required for tissue classification over the wound is a difficult task. As direct classification on the pixels with simple thresholds on separate color components proved to be inefficient due to the variability and non homogeneity of tissues, spatial continuity has been searched for through a segmentation process. The classification process is thus driven by the segmentation one [10]. During the following step, a tissue database has to be built, covering all wound grades, locations and healing status. This recording task is performed mainly by the care staff, under the supervision of clinicians, as the wounds are only visible after cleaning during dressings. In many studies, the tissue samples were manually extracted by asking the clinician to pick square homogeneous regions of interest inside the images [18]. Complete classification would require that during the tests these samples be automatically located in the image without a priori knowledge. Moreover, using photos instead of examining the patient has an impact on expert performance [14]. When several experts are consulted, it appears that inter-observer variability is important, exceeding intra-observer variability [5]. For classification purposes, several region descriptors have been widely tested on wound tissue samples. Color histograms are typical of such statistical data gathering as tissue descriptors [40], [7], but it has been proved that tissues could still not be classified robustly in a large collection of images. The best results were obtained when the classification was limited to two types of tissue [37] or by multiplying the tissue classes using hybrid classes [51]. To improve the results, texture parameters have finally been added [18], [13]. Several classification tools have also been addressed and supervised ones have been shown to surpass the others [6], [34]. In many studies, assessment of the tissue type is not achieved (only the grade or the surface of the wound is provided) [26], [43], [20] or only partially.

Beside the methods of tissue classification, some prototypes based on structured light techniques [36], [27], [31] or photogrammetry [8], [32], [40] have been presented in order to obtain spatial measurements. The prototypes called MED-PHOS (Medical Digital PHotogrammetric System) [32], [33] and MAVIS (Measurement of Area and Volume Instrument System) [40] are typical of this approach. However, these cumbersome and complex systems are unsuited to clinical practice which requires a low cost, handy and simple tool operated by a nurse. Portable industrial 3D scanners have also become available and have been experimented on wound measurement. Examples are the Minolta V series [42], [9] and 910 series [16] and the Fastscan Polhemus system [29], all based on a laser stripe scanning, the latter being moreover a hand-held system due to magnetic sensing of the system pose. More recently, a commercial system has been proposed by Aranz Medical company (*http://www.aranzmedical.com*): the SilhouetteMobile system is based on a personal digital assistant (PDA) equipped with a small digital camera with embedded laser lighting. This tool is very expensive compared to a simple digital camera and tissue classification is not supported. Wound extraction must be done by drawing around the wound outline using a stylus on the PDA screen.

The systems that are most closely related to our work are MAVIS-II [24] and the Derma project [42], [9]. MAVIS-II uses a reflex digital camera equipped with special dual lens optics to record two half images from slightly different viewpoints. This system competes with some 3D laser scanners, but it suffers from several drawbacks;firstly, a costly digital reflex camera is required to adapt the special dual lens;secondly, the stereo lens adapter constrains both the field of view and the focusing distance and consequently the size of imaged wounds, as it must be located around the intersection of a dual light point projector also required;finally, the tissue classification problem is not addressed. In the Derma project, wound measurements and tissue classification have both been tackled, as it enables shape and natural texture grabbing, but the classification process remains user-assisted as seeds need to be manually pointed inside the wound and a similarity measure adjusted to control the merging process. Moreover, it is based on a very costly Minolta 3D scanner, putting it beyond the reach of routine clinical practice. In sum, wound assessment using imaging techniques is still mainly based on 2D data processing and no existing system combines color tissue classification and wound 3D shape reconstruction into a single low-cost tool for enhanced and complete wound assessment.

### C. Proposed framework

The main objective of this work is to present a complete and user-friendly tool that can be easily implemented by care staff, in order to automatically monitor the healing process. Such a tool can provide efficient measurements necessary both in clinical practice to monitor the progress of healing, and in research to evaluate new treatment. In contrast to the aforementioned methods dealing separately with wound shape and tissue classification tasks, we suggest going further toward a very low-cost and easy-to-use device using a simple standard digital camera. This is the only way to ensure that quantitative wound assessment will be widely practised by clinical staff, who have to handle high rates of pressure sore prevalence in most hospital departments. A complete wound assessment tool has therefore been developed, combining a sharp 3D model of the wound and a robust classification tool for color tissues

(Fig. 1). The part concerning computation of a 3D model has been presented in refs. [2], [3], [45], while the classification tool has also already been described in detail in [47], [48], [50]. Our strategy, introduced in ref. [49], involves combining several viewpoints both to improve the classification process itself, as tissue aspect is view-dependent, and also to compute real surfaces by mapping the regions onto the 3D model. In this paper, we demonstrate that the integration of these two modules gives access to enhanced tissue classification and measurements, as several views are combined to classify the tissues and the results mapped directly onto the mesh surface to obtain real tissue areas. Only a brief presentation of the classification method and classifier design will be given here, to focus particularly on the merging of single view classification results, called here multi-view classification.

The paper is organized as follows: Section II details the wound image processing chain involved in our application, in particular, color correction and image segmentation and labeling by experts; Section III presents the selection of tissue descriptors and the tissue classifier design and testing;Section IV details our multi-view strategy for enhanced classification, in particular the mapping and merging of the classification results on the 3D model;experimental results are presented in Section V to address the repeatability and the robustness of the multi-view classification, as well as access to metric assessment with the 3D model. The last section presents our conclusions and the main perspectives of this work.

## II. WOUND IMAGE PROCESSING

### A. Wound database building and color correcting

A database of several hundreds of color images (3 Mpixels, 24 bits) of chronic wounds was constituted with the help of clinical staff in several hospitals, in order to obtain a representative set of images for different tissue types as input for the wound image processing chain (Fig. 1). A preliminary requirement was to acquire good quality wound images covering all types of pathologies. A variety of chronic wounds (sacrum and heel bed sores, leg ulcers, diabetic feet, etc.) was collected from different care services including gerontology, rehabilitation, diabetology and dermatology. Other types of wound images such as traumatic, post-operative images can be collected from surgery services, but neither these particular types nor burn scars where included in this study. The images were taken with different digital cameras under uncontrolled lighting conditions, following a specific protocol that included several viewing angles for each single wound. Special attention was paid to color constancy to enable reliable tissue classification. Practically, a standard 24 patch Macbeth color checker pattern was introduced in the field of view, providing also the scale factor in the different images. It addressed two distinct tasks: firstly, obtaining a constant response from the digital camera by identifying lighting conditions and secondly, calibrating the digital camera color response. After correction, colorimetric stability reaches the limit of discernibility for a human observer, as $CIE \Delta E_{ab}$ was reduced to $3.5 \pm 2.9$. This step is essential in a telecare environment involving several care centers where the practical conditions differ widely.

### B. Multi-expert image labeling

A selection of fifty typical wound images was extracted from the database and submitted to a group of clinicians in order to label them according to the classical color code using our graphical interface. This interface enables the user to draw closed boundaries using a pen tablet or a mouse and to fill in the area with color labels. In this medical imaging application, it provides the input data for building the absolute medical reference by merging expert data, as the clinician tracings are freeform. Precautions were taken concerning the lighting conditions and the color calibration of the graphical screens to ensure faithful rendering during the labeling sessions. Patients corresponding to the wounds had not been previously seen by the clinicians, in order to eliminate any external influence. In any case, the considerable distances between care centers prevent the clinicians from visiting all the patients, precluding any merging of the deduced labels. Tissue labeling was carried out twice, one month apart, by four clinicians, in order to measure the accuracy during labeling. We used the normalized Overlap Score (OS), classically used to compare a segmentation S to a ground truth G by $|S \bigcap G|/|S \bigcup G|$. It was averaged over all the tissue samples weighted by their corresponding areas. As $0 \leq OS \leq 1$ it can be expressed as a percentage. The OS obtained by the same clinician labeling, one month apart, are given in Table I. The labels of the different clinicians were merged to obtain a single reliable medical reference for evaluation of the algorithm. A majority vote criterion was applied to label each pixel and only the pixels with a confidence level greater than or equal to 75% were retained. The OS were then computed between the clinicians' tracings and this medical reference (Table II).

| Tissue class | Expert OS (%) | | | | |
|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **Average** |
| Granulation | 73.9 | 65.0 | 55.9 | 66.7 | 65.4 |
| Slough | 77.4 | 73.8 | **55.3** | 71.4 | 69.5 |
| Necrosis | 81.4 | 91.4 | 59.5 | **81.9** | 78.6 |
| **Expert average** | 77.6 | 76.7 | 56.9 | 73.3 | **71.2** |

TABLE I
INTRA EXPERT OVERLAP SCORES.

| Tissue class | Expert OS (%) | | | | |
|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **Average** |
| Granulation | 73.2 | 72.7 | **39.3** | 71.0 | 64.1 |
| Slough | 58.9 | 76.0 | 42.8 | 70.8 | 62.1 |
| Necrosis | 79.3 | 77.1 | 82.6 | **85.6** | 81.2 |
| **Expert average** | 70.5 | 75.3 | 54.9 | 75.8 | **69.1** |

TABLE II
OVERLAP SCORES BETWEEN EXPERTS AND MEDICAL REFERENCE.

The results obtained call for several comments. To begin with, the differences in Tab. I may be explained by variable levels with respect to both computer use and wound classification. Another reason resides in variations of screen color rendering: the screens had been calibrated but the ambient lighting conditions and the tuning of screen brightness and contrast may have changed between the two labeling sessions.
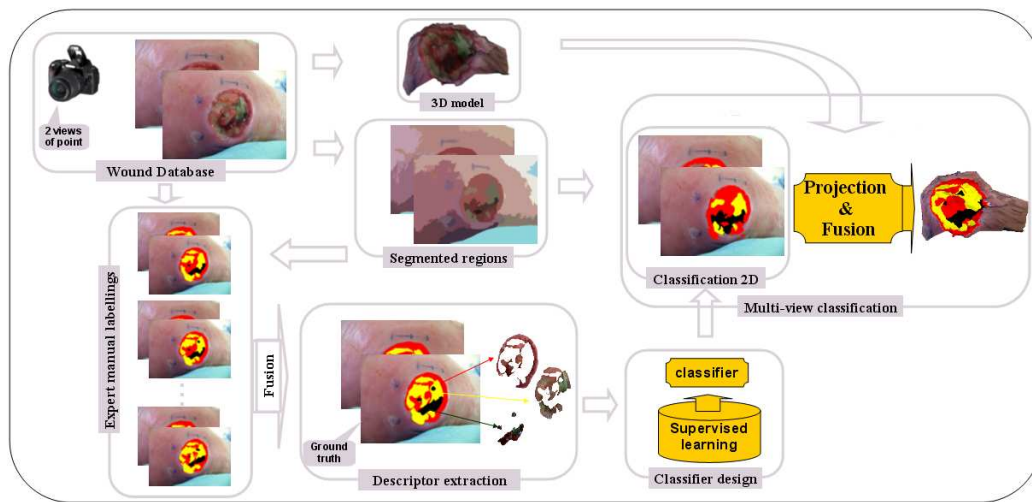
Fig. 1. Complete wound assessment tool combining 3D measurement and tissue classification capabilities.

The OS between clinicians and the reference remain moderate (55% to 76%) as the labels rely on the subjective diagnosis of the clinician, who may not be familiar with wound assessment on a computer, without direct patient examination. When the patients have been seen by the clinicians, the inter-expert OS are improved, as the clinicians can inspect the wounds from many point of view and rely on their knowledge of the patients. Consequently, image-based wound assessment should assist but not replace the clinician in the decision process. To finish, the scores obtained for the slough tissue class are the lowest, as this type of tissue is often mixed with necrosis and because granulation grows under slough areas; moreover, none of the clinicians produce exactly similar labels one month later (57% to 78%). A very practical issue of these inter- and intra-expert OS is to provide a performance target to be reached by the automatic classification process.

*C. Automatic wound segmentation*

As direct pixel classification proved to be inefficient, we tested several advanced unsupervised segmentation algorithms efficient on textured images, to provide an automatic delineation of tissue samples and to simplify the following classification step by extracting more robust and discriminating local attributes on tissue areas: the Color Structure Code (CSC) [41], Efficient Graph-Based Image Segmentation (EGBIS) [17], Mean Shift [12] and J-SEG [15]. The best results were obtained with J-SEG (average OS of 73.1%), by assigning to each unlabeled region the class of tissue mainly represented in the corresponding area of the medical reference and computing the OS between the assigned labeling and the medical reference. The high OS confirm that segmentation is a valuable pre-processing step before classification for an automatic delineation of tissue areas. The power of the J-SEG algorithm resides in the separation of the segmentation process into two independent processing stages: color quantization and spatial segmentation. These two steps are controlled by free parameters which must be precisely tuned to adjust the number of output regions: sub-segmentation leads to misclassification of regions which include several classes of tissues, while over-

segmentation increases the subsequent processing time and results in classification errors on tiny regions where reliable and robust descriptors cannot be obtained without statistical evidence.

As the OS between clinicians' manual tracing and pre-segmented regions are quite relevant, a second interface was designed to allow clinicians to directly label pre-segmented regions automatically (Fig. 2). This interface is far less demanding for the experts as only simple mouse clicks on the regions are required. On the same basis as in Sec. II.B, a practical medical reference was developed by merging the labeling results of several experts. A total of 850 significant tissue regions was selected among the 1200 samples extracted from the fifty typical wound images previously submitted to the experts. The samples selected exhibit the usual types of tissues (48% Granulation, 38% Slough, 14% Necrosis, relative to the number of samples), with at least 75% agreement between the experts. We observed that the OS performance in tissue labeling is improved when using automatic wound pre-segmentation. This increase is artificial, because the expert no longer controls the delineation process (he has to label each region with a single label). One important point is that the training of the classifier presented below is improved when using a ground truth built on labeled regions instead of tracings. The reason for this improvement may be that in the first case, the classifier is trained on similar regions (obtained with the same segmentation algorithm) during the training and test steps.
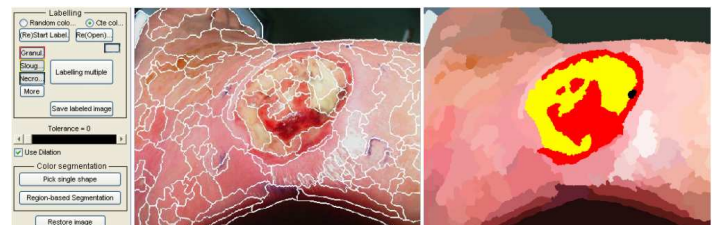


Fig. 2. Graphical interface for the expert manual labeling.

## III. Tissue classification from a single view

### A. Feature vector generation

We focused on the color and texture descriptors already applied in dermatological imaging systems, especially for wound and ulcer tissues. Color is probably the most dominant image cue as demonstrated by the red/yellow/black healing visual assessment during clinical visits. The color descriptors we have extracted are: the Mean Color Descriptor (*MCD*), the locally adapted Dominant Color Descriptors (*DCD*) based on the Mean Shift iterative color clustering algorithm [12], and 2D and 3D color histograms [7] [25] [18] tested in different color spaces and sizes. Texture is also a relevant cue for tissue description if it is computed in a neighborhood properly scaled to the local structure. While clinical investigations indicate that it is less relevant than color, it nevertheless provides complementary data. Several well-known texture descriptors from color components (RGB) or gray level image (GL) were extracted from the wound tissue regions. These descriptors are: Gabor based features (*GAB*) calculated on five scales, Local Binary Pattern (LBP) histograms [25], Haralick Gray Level Co-occurrence Matrix features (GLCM) [21] and the normalized texture contrast and anisotropy (*CA*) (Tab III). The computation time stands for a $512 \times 384$ image processed by a PC Pentium 4 CPU 3.4 GHz. It may vary slightly, depending on the number of regions and the actual size of the wound inside the image. To provide more consistent evaluation and efficiency of the image processing chain, the performance of these descriptors was also evaluated directly at the output of the classifier (see Sec. III.C).

| Descriptor | Symbol | Space | Size | Time (*sec*) |
|---|---|---|---|---|
| Histogram | *h-RGB* | 3D-RGB | 64 | 0.6 |
| Histogram | *h-LAB* | 3D-Lab | 218 | 1.56 |
| Histogram | *h-HSV* | 2D-HS | 256 | 1.99 |
| Histogram | *h-rg* | 2D-rg | 256 | 1.60 |
| Mean-Dominant | *MCD-DCD* | Lab/RGB | 8 | 0.41 |
| Local Binary Pattern | *h-LBP* | GL | 256 | 5.3 |
| Local Binary Pattern | *h-LBP* | GL | 59 | 1.5 |
| Co-occurrence matrix | *GLCM* | RGB | 15 | 0.48 |
| Gabor filter | *GAB* | RGB | 15 | 1.6 |
| Contrast-Anisotropy | *CA* | GL | 2 | 0.23 |

TABLE III
COLOR AND TEXTURE DESCRIPTORS.

### B. SVM Classifier design

Initially designed for binary classification after supervised learning, Support Vector Machines (SVM) are also used for multi-class problems through one against all and one against one strategies [22]. Non linear class separation in low dimension space may result in smart separation in higher dimension space, using a suitable kernel function. The key point of the SVM classifier design remains the choice of the kernel function, as this depends on the image database and input descriptors [11] since no universal kernel will fit all applications. The SVM classifier selected here is a soft-margin algorithm (so-called C-SVM) available online at *http://www.csie.ntu.edu/ucjlin/libsvm*. It has been tested by

computing ROC curves for several classical kernels: linear, polynomial, radial basic function (RBF) and perceptron (Fig. 3). After the selection of a particular kernel, its parameters must be tuned. In the case of the perceptron kernel finally selected, there is only one free parameter; it controls the penalty of the classification error and has been adjusted by a line search technique.
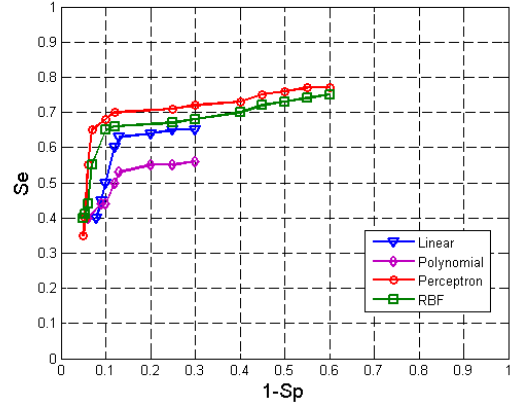


Fig. 3.   ROC curves obtained by four different kernels

During the classifier learning stage, the subset of 850 significant region samples labeled by the experts was divided equally into training and test subsets, then several iterations were applied through a k-fold cross validation with k=5 to randomly select the training set before averaging the results. Each feature vector was tested with different settings of the classifier parameters during a backward and forward process. To evaluate the quality of a diagnostic method, clinicians usually compute predictive measures such as sensitivity (Se), specificity (Sp), success rate (Sr) and overall accuracy (Oa), which are reliable performance estimators [1]. These estimators enabled us to compare the discriminating power of the input descriptors (Tab. IV).

| Feature | Descriptor | Se (%) | Sp (%) | Sr (%) | Oa (%) |
|---|---|---|---|---|---|
| | *h-RGB* | 58 | 87 | 72 | 80 |
| | *h-LAB* | 66 | 87 | 76 | 82 |
| **Color** | *h-HSV* | 62 | 87 | 75 | 81 |
| | *h-rg* | 57 | 86 | 72 | 80 |
| | ***MCD-DCD*** | **67** | **89** | **78** | **84** |
| | *h-LBP* | 30 | 78 | 54 | 66 |
| | *m-LBP* | 29 | 77 | 53 | 66 |
| **Texture** | *GLCM* | **54** | **82** | **68** | **72** |
| | *GAB* | 47 | 81 | 64 | 71 |
| | *CA* | 32 | 79 | 55 | 68 |
| **Col.+Tex.** | ***MCD-DCD+GLCM*** | **77** | **92** | **84** | **88** |

TABLE IV
PREDICTIVE POWER OF SEVERAL COLOR AND TEXTURE DESCRIPTORS.

The numerous tests involving several data sets show that the best results are finally obtained by combining 22 attributes

including MCD-DCD as color descriptor and GLCM as texture descriptor as input of the C-SVM classifier implemented with a perceptron kernel (parameter C = 100). To evaluate the performance of the SVM approach more completely, we also compared it to other strategies, with the same descriptors as input. We selected other classical supervised methods, such as K-NN and Fuzzy K-NN and k-Means as an unsupervised one. As before, the same data set was randomly divided into learning and test sets to program the classifier (Table V).

| Classifier | $Se$ (%) | $Sp$ (%) | $Sr$ (%) | $Oa$ (%) |
|---|---|---|---|---|
| K-NN | 63 | 86 | 75 | 80 |
| Fuzzy K-NN | 66 | 87 | 77 | 81 |
| K-Means | 39 | 77 | 58 | 68 |
| SVM | **77** | **92** | **84** | **88** |

TABLE V
COMPARISON OF SEVERAL CLASSIFIER PERFORMANCES.

It can be seen that unsupervised learning is inefficient in classifying the tissue samples into three different classes, whereas the supervised approach appears to be quite relevant for this kind of problem and thanks to its generalization ability, the discriminating power of SVM is unequalled.

### C. Classifier testing

To validate this classification tool applied to chronic wound assessment, global statistical scores are not sufficient. To come closer to clinical measurements, two series of tests are required.

In the first one, we examine separately the classifier performance in relation to each class of tissue by counting well classified and misclassified regions. Ten iterations were again applied to randomly select the training set and the final results were averaged. To evaluate the degree of agreement between two judgments, the Kappa statistical coefficient is widely used in the medical field, as it avoids nonsignificant high scores when randomly voting [28]. This statistical indicator, varying between 0 and 1, can be calculated by $K = \frac{P_o - P_e}{1 - P_e}$ where $P_o$ is the relative observed agreement and $P_e$ the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly voting for each category.

The results of Tab. VI demonstrate the efficiency of the classifier against individual experts compared to the medical reference. Furthermore, it produces high scores on necrotic regions, which are critical in medical decisions, as immediate and invasive clinical treatment may be required if such regions are detected. An average Kappa coefficient of 0.81 was obtained, which is considered as quite good agreement.

In the previous tests, the classification performance was examined globally for each type of tissue by counting the number of well classified regions. However, in clinical applications, the misclassification of a small area does not have the same impact as that of a large one. A second series of tests was therefore carried out in which the quality of assessment is measured directly on the same set of fifty typical wounds, with area weighted scores (Table VII).

| Tissue class | Experts | | | | |
|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **SVM** |
| Granulation | **0.87** | 0.74 | 0.68 | 0.82 | 0.80 |
| Slough | 0.75 | 0.78 | 0.54 | **0.80** | 0.79 |
| Necrosis | 0.65 | 0.87 | 0.44 | 0.79 | **0.85** |
| **Average** | 0.75 | 0.79 | 0.55 | 0.80 | **0.81** |

TABLE VI
KAPPA COEFFICIENT FOR TISSUE SAMPLE CLASSIFICATION.

| Tissue Class | Experts | Algorithm |
|---|---|---|
| Granulation | 64.1 | 85 |
| Slough | 62.1 | 70 |
| Necrosis | 81.1 | 83 |
| **Average** | **69.1** | **79.3** |

TABLE VII
AVERAGED OVERLAP SCORES OF EXPERTS AND CLASSIFIER IN (%)
COMPARED TO THE MEDICAL REFERENCE.

After averaging on the college of experts, it appears that the classifier performs better than a single expert. This can be attributed to the fact that the learning step relies on a reference established from several experts, and is thus based on stronger evidence. Fig. 4 shows the classification results obtained on different wound images. Several features can be pointed out. Firstly, the power of separation of tissue classes is constant over various locations and pathologies (e.g. sacrum pressure sore, leg ulcer and ankle wound). Secondly, it is not affected by the skin tones of the different patients or the lighting conditions encountered. Moreover, the classifier is not confused by reflections on moist granulation regions.
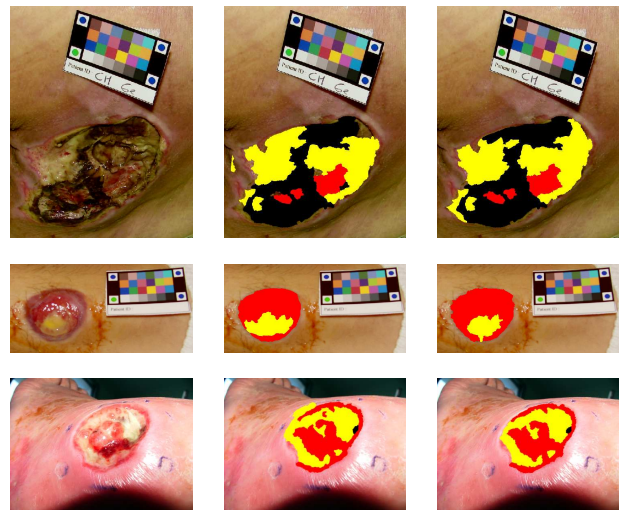


Fig. 4. Three examples of automatic wound assessment compared to the medical reference. ($1^{st}$ column) original image ($2^{nd}$ column) automatic classification ($3^{rd}$ column) medical reference.

However, these 2D classification results obtained by single view assessment suffer from a lack of repeatability between the visits of patients. The reason for that is simple: during the follow-up, pictures are taken free hand by the nurses and under uncontrolled lighting. It has been established that a deviation of $20°$ of the optical axis from the normal of the wound typically leads to an underestimation of surface of around

10% [38]. This is due to lighting variations which modify the colors and perspective effects from distant viewpoints, inducing significant bias in the classification results and preventing real surfaces from being computed. In the following section we present our approach to multi-view classification. It takes advantage of the geometric correspondence of the triangular zones of the wound in two different views provided by the 3D model, to achieve a more robust tissue classification, referred to as multi-view.

## IV. MULTI-VIEW CLASSIFICATION

Like the clinician, who draws on many observation points to provide a reliable diagnosis, a multi-view technique should allow more robust results. We therefore propose to use the dimensional information captured from a multi-view model because reliable wound assessment must provide reproducible results, regardless of the position and orientation of the camera. Based on the 3D reconstruction of the wound model [45], the main idea is to combine the color information of the regions, the calculation of points of view and the relief in order to achieve a more robust classification and also access to real surfaces.

### A. Mapping of the classification results on the 3D wound model

Clinicians establish their diagnosis visually on the photographed wound, with the help of a red-yellow-black scale placed in the camera field, corresponding to the three types of tissue. However, this diagnosis is also based on their observations of the wound during examination of the patient. The clinician's assessment of tissues can thus be seen as a combination of colorimetric information (image plane) with shape information (through observation with the human eye). To get closer to the process of patient examination, it was therefore natural to take wound images from different points of view. A selection of ten wounds imaged from at least three different viewpoints was therefore made, in order to produce several input data sets for the 3D reconstruction process. A 3D model was obtained from these images using uncalibrated vision techniques completed by original refinements to obtain semi-dense matching between widely separated views. Typically, 3000 to 4500 matches are obtained in $1024 \times 768$ image pairs, making it possible to match homologous regions in each view and to merge classification results [4].

We can illustrate the dependence of the classification assessment on the point of view by a simple back projection on a 3D model computed from two views of a wound. To do this, the classification result from each of the single views was mapped onto the 3D model separately in order to label the triangular mesh. Each triangle is labeled according to its tissue highest score and then the surface of each type of tissue can be computed by summing the triangles belonging to the same class. Fig. 5 shows the variation of the cartography mapping on a 3D ulcer model according to the classification results obtained with the single view approach.

The 3D model allows accurate comparison of single-view classifications since the differences are expressed in $cm^2$

and not in pixels. Table VIII presents the area of tissue surfaces calculated in $cm^2$ in each view and the corresponding variations expressed as a percentage of the total wound area, when mapping single view classification on the 3D models presented in Fig. 5.
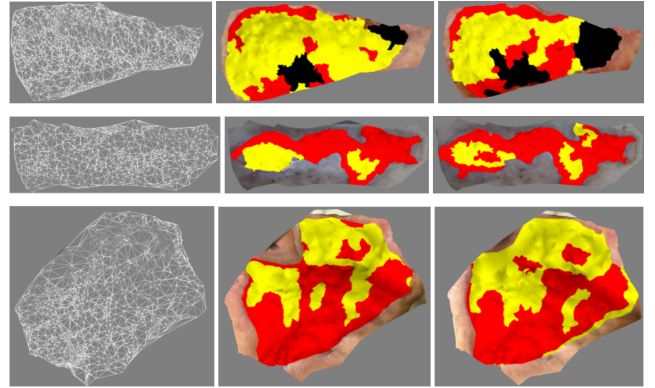


Fig. 5. Influence of the point of view on the classification results. ($1^{st}$ row) deep sacrum pressure sore ($2^{nd}$ row) leg ulcer ($3^{rd}$ row) diabetic foot. The 3D model ($1^{st}$ column) is labeled with the the classification results from the first view ($2^{nd}$ column) or the second view ($3^{rd}$ column).

| Tissue class | Example 1: deep sacrum pressure sore | | |
|---|---|---|---|
| | view 1 ($cm^2$) | view 2 ($cm^2$) | variation (%) |
| Granulation | 12.9 | 18.2 | **9.8** |
| Slough | 36.9 | 27.4 | **17.6** |
| Necrosis | 4.2 | 8.4 | **7.8** |
| | Example 2: leg ulcer | | |
| | view 1 ($cm^2$) | view 2 ($cm^2$) | variation (%) |
| Granulation | 39.0 | 42.1 | **6.6** |
| Slough | 6.0 | 4.9 | **2.3** |
| | Example 3: diabetic foot | | |
| | view 1 ($cm^2$) | view 2 ($cm^2$) | variation (%) |
| Granulation | 21.4 | 20.6 | **1.8** |
| Slough | 18.9 | 25.1 | **13.6** |

TABLE VIII
MAPPING OF SINGLE IMAGE CLASSIFICATIONS ON THE 3D WOUND MODEL

The differences obtained reflect the effect of perspective projection in the image and the relief of skin ulcers. This experiment confirms the limitations of the single-view approach for patient follow-up and the advantage to take into account 3D information. Therefore, the results of tissue labeling coming from each image need to be merged on the wound 3D model to achieve accurate and robust assessment.

### B. Merging algorithm

Because of the epipolar geometry, the neighborhoods corresponding to a single facet of the 3D discrete model can be linked in the images, making it possible to establish the link

between two single view tissue labels using the 3D model of the wound. In previous related work, the analysis was limited to single-view (normal to the capping plane of the wound) since a model of 3D anatomical wounds was not available.

The classification results obtained for each of two images of the wound, taken from different viewpoints, have to be merged based on the 3D model obtained from these two views. The 3D model, composed of a mesh of triangular facets, is projected onto the stereo pair to provide a 2D Delaunay mesh of triangles in each image. Due to the point correspondences between the two images, each triangle in the left image has a homologous one in the right image. As they cover the same portion of the wound but do not belong necessarily to the same class (Fig. 6), it is necessary to merge the results of tissue labeling coming from each image.
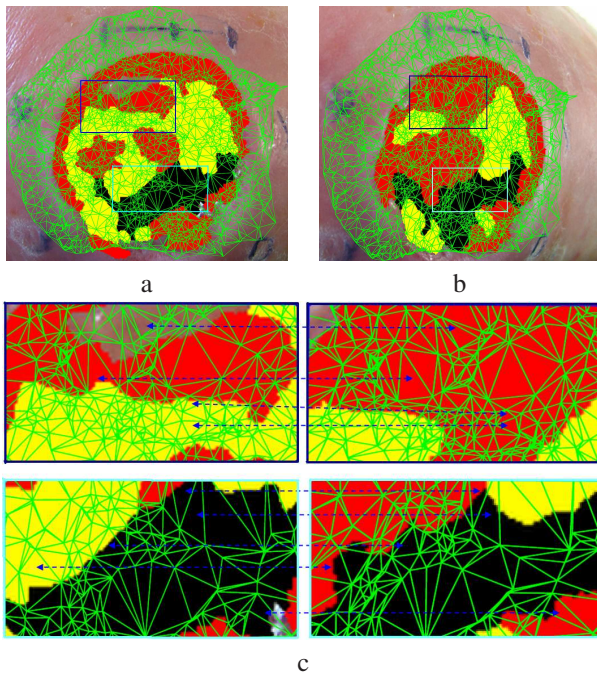


Fig. 6. Projection of the 3D mesh on the classification results of two views. (a-b) projected model on the classification results of the two images (c) two homologous selected zones, illustrating the correspondence between homologous triangles coming from the same facet of the model.

The merging strategy we have experimented is summarized by the following algorithm (Fig. 7). For each facet of the 3D model, the proportion of each type of tissue is calculated in each of the homologous triangles, according to its red/yellow/black color (standing for granulation/slough/necrosis tissue). In the obvious case where the homologous triangles contain only one type of tissue identically classified in both images, the facet of the 3D model is labeled according to this type. Otherwise, the triangle is recursively split along the median line of its longest side into two sub-triangles. This splitting process is stopped when the area of the facet reaches a minimal value (cutting threshold).

The 3D points constituting the triangular mesh result from a matching process between the points of interest detected in the image and are not equally distributed on the wound surface but are highly concentrated in textured regions and reciprocally

very sparse in homogeneous regions. Consequently, the typical size of the triangles may differ across the model surface. At this stage, if the splitting of the triangle does not provide homogeneous sub-triangles, three possibilities must be tested:

- the first is to apply the dominant class criterion and assign this dominant class to the facet of the 3D model, when it is common to both homologous triangles.

- otherwise, i.e. when the dominant classes are different in the two images, we compute the two solid angles carried by the facet and resulting from the optical center of the camera in each view. If the difference between the two angles is significant, this means that the facet is viewed diagonally in one view and from a more frontal angle in the other. In this case, the facet model is labeled entirely by the class corresponding to the largest solid angle. This argument is based on the fact that the classification process is more efficient when the triangle is close and pointing to the camera optical center.

- in the last case, when the difference between the solid angles is not significant, a confidence level must be computed to estimate in which image the classification is likely to be the most reliable. It is based on the class probability of belonging to a given class of the two regions containing the homologous triangles within single-view classification. The triangle is labeled with the label of the region classified with the highest probability at the SVM classifier output.

```
Require: classification results on two views + 3D model
  Result: multi view classification
  project the 3D triangular model on each view
  for all triangles of the 3D model do
    find tissue percentage on each view;
    if the two views have the same tissue class then
      label triangle with common tissue class;
    else
      compute 3D triangle surface;
      if surface > threshold then
        split the triangle & go back to for loop;
      else
        identify dominant tissue in each view
        if same majority class then
          label triangle with this majority class;
        else
          compute the solid angle in each view (S1, S2)
          if S1 ≫ S2 then
            label triangle corresponding to 1st view;
          else if S2 ≫ S1 then
            label triangle corresponding to 2nd view;
          else
            compute the class probabilities in each view
            label triangle by max probability class;
          end if
        end if
      end if
    end if
  end for
```

Fig. 7. Merging algorithm.

## C. Statistical analysis of the merging process

When all the triangles have been scanned, the label resulting from the fusion is mapped onto the 3D mesh and the real tissue surfaces are computed on it. Clearly, single-view classification could not give access to the real surfaces, as only pixel counting on areas projected in the image plane was possible.

Obviously, the fusion of single view classification results may be applied to the wounds labeled by the experts. To do this, we established for each pair of images a 3D medical reference by merging the two 2D medical references coming from the left and right images and mapping them onto the 3D model. The process of creating the latter relies on the same fusion strategy presented in Algorithm 1, but without taking into account the criterion of class probability estimates since no classifier is used here. It could be replaced by the level of confidence obtained from several manual expert labelings. This 3D medical reference is used to evaluate the improvement due to the fusion of 2D classifications.

To analyze the management of the triangle labeling process through the fusion algorithm, we perform it on fifteen pairs of wound images labeled by both manual labeling (expert) and automatic classification (classifier). The histogram in Fig. 8 shows the total surfaces of the triangles and their number, labeled at each step of the fusion algorithm applied on the 3D models.
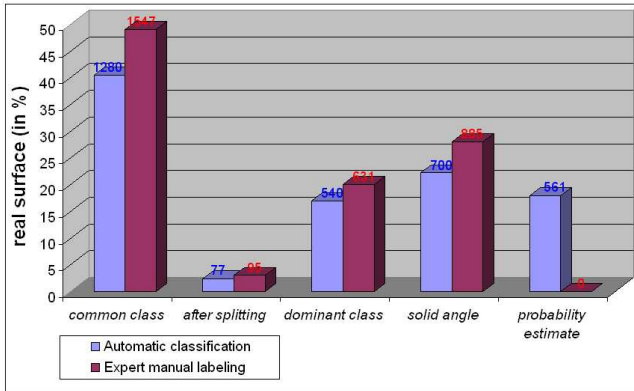


Fig. 8. Distribution of the labeling in the fusion algorithm.

We can derive from this histogram that for more than half of the model surface, the classification results are strongly dependent on the viewpoint, as only 40% of the total surface was labeled with the same class in both views. Only a few triangles are concerned by the splitting step; this is because of the semi-dense 3D model in our matching process. However, about 20% of the wound model area is labeled according to solid angle criteria. Finally, for about 35% of the model surface, the two criteria of dominant class and probability estimates need to be computed. In this way, we highlighted qualitatively the dependence of the single-view classification and the need to combine several views to ensure successful classification. Note that for the manual labeling, the percentage of tissue surface labeled by the probability criterion is zero because neither the degree of confidence of the expert labels nor the degree of expertise for each expert were taken here into consideration.

One important observation is the strong similarity between the histograms corresponding to manual labeling and automatic classification. It shows clearly that both the expert and the algorithm produce view dependent results. In particular, the common class category receives similar scores, with a slightly higher one for the experts. This difference may be explained by the fact that the expert knows that the different viewpoints are concerned with the same wound and he tries naturally to preserve as much as possible the spatial coherence between his labelings.

## V. Experimental results

### A. Repeatability of the classification

The manual selection of picture viewpoints and the random nature of the 3D reconstruction algorithm may impair the repeatability of the classification results. Since the matching strategy in the reconstruction stage is developed on the semi-dense matching algorithm incorporating a RANSAC step. we never obtain two strictly identical models in successive executions of the algorithm [45]. For this reason, several experiments were done to evaluate the repeatability and stability of multi-view classification results, in relation to viewpoint changes and 3D reconstruction differences. Practically, the stability of the process was evaluated firstly on different 3D models calculated from the same pair but for several reconstructions (Fig. 9).
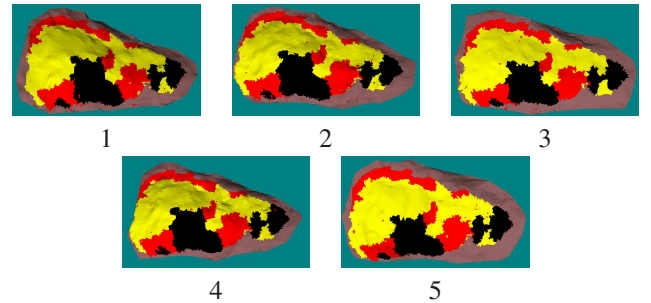


Fig. 9. Stability of the process over five reconstructions of the same pair of images. The multi-view labeling (3D) is mapped onto five 3D models computed successively from the same pair of images.

Secondly, the repeatability of the classification was evaluated on different pairs of images of the same wound. For each wound tested (Fig. 10 - Fig. 11 and Fig. 12), three images of the wound were taken from different viewpoints (one frontal view and two side views), constituting two stereo pairs. A 3D wound model was reconstructed using each stereo pair and the multi-view classification approach was applied on it. 3D model reconstruction was then executed five times on each stereo pair to provide ten models of each wound onto which the tissue labels were mapped.

It appears that the classification results are quite stable even with the slight changes in 3D model geometry at each new model reconstruction (Table IX). The influence of the 3D reconstruction on the repeatability is greater than that of the view selection but the latter is easily reduced by the simple averaging of several executions of the reconstruction algorithm. In particular situations, this repeatability can be
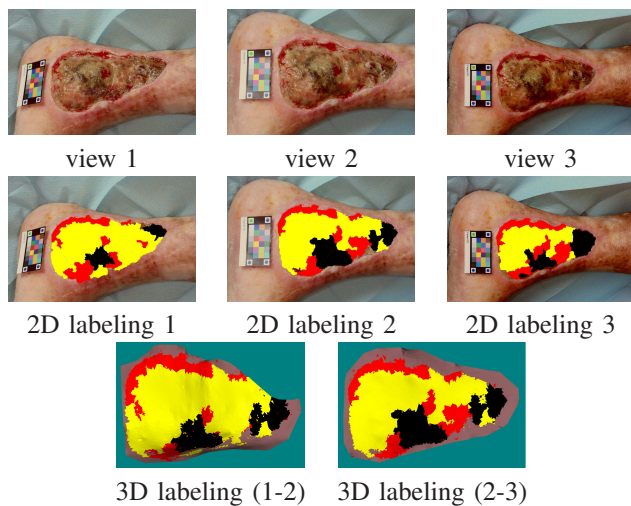
Fig. 10. Repeatability of the classification over five reconstructions for two pairs of images of a wound (leg ulcer).



Fig. 12. Repeatability of the classification over five reconstructions for two pairs of images of a wound (heel pressure sore)
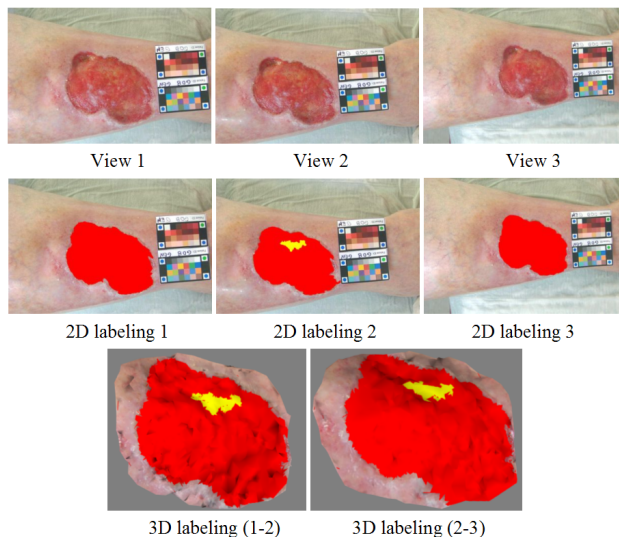


Fig. 11. Repeatability of the classification over five reconstructions for two pairs of images of a wound (ankle ulcer)

degraded: this is the case when the images are captured along an orbital path because the self-calibration algorithm becomes inefficient;this is again the case in poor lighting conditions resulting in a sparse 3D model with shape distortions. Such situations are avoided by a strict procedure during snap-shots, to be followed by the clinicians.

### B. Robustness of the classification

Considering the merging algorithm, it is common sense to consider that combining classifications from several views should improve the results, as it enriches the knowledge available in single-view. However, this needs to be tested on our wound image database to assess quantitatively the real advances in this medical application. To evaluate the improvement due to the fusion of 2D classifications, we have
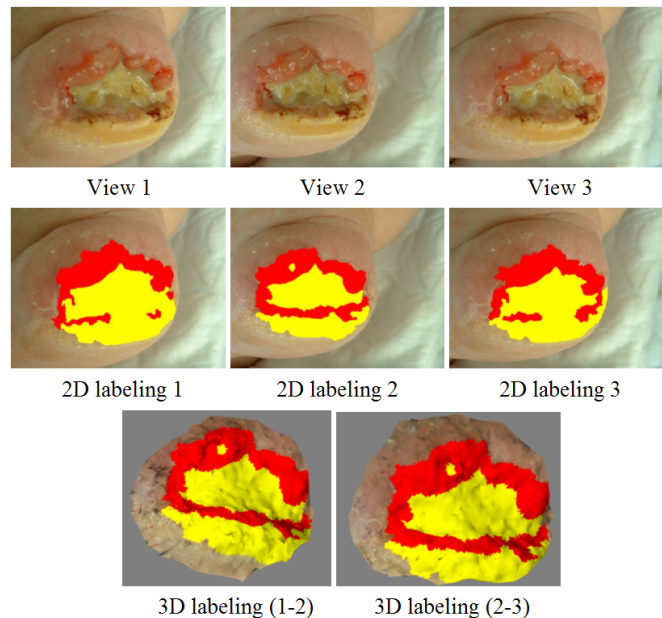
| Tissue class (in % ) | Example 1:leg ulcer | |
|---|---|---|
| | **1-2** | **2-3** |
| Granulation | 17.4±1.8 | 17.6±0.5 |
| Slough | 67.0±1.4 | 67.6±0.5 |
| Necrosis | 15.6±0.3 | 14.6±0.5 |
| | Example 2:ankle ulcer | |
| | **1-2** | **2-3** |
| Granulation | 96.3±0.6 | 96.5±0.7 |
| Slough | 3.7±1.4 | 3.5±0.5 |
| | Example 3:heel pressure sore | |
| | **1-2** | **2-3** |
| Granulation | 38.8±0.8 | 38.5±1.7 |
| Slough | 61.2±0.8 | 61.5±1.7 |

TABLE IX
REPEATABILITY OF THE CLASSIFICATION OVER FIVE RECONSTRUCTIONS
FOR TWO PAIRS OF IMAGES OF WOUND IMAGES

to compute the overlap scores obtained for single-view (2D) and multi-view (3D) classifications. Finally, we compared the scores between the 2D medical reference and 2D automatic classification on the one hand, and the overlap scores between the 3D medical reference and 3D classification results on the other hand (Table X). Therefore we also calculated the Kappa coefficient in the 2D approach (between classifier and 2D medical reference) and the 3D approach (between 3D classification and 3D medical reference) (Table X).

| Tissue class | Overlap score (%) | | Kappa coefficient | |
|---|---|---|---|---|
| | 2D | **3D** | 2D | **3D** |
| Granulation | 79.8 | 81.4 | 0.82 | 0.84 |
| Slough | 69.3 | 72.0 | 0.75 | 0.77 |
| Necrosis | 60.7 | 67.9 | 0.73 | 0.77 |
| Average | 69.9 | **73.8** | 0.77 | **0.79** |

TABLE X
2D/3D OVERLAP SCORES AND KAPPA COEFFICIENTS AVERAGED OVER
THE WOUND DATABASE.

It can be seen that the agreement between medical reference and automatic classification is globally improved after the merging step. The improved performance of multi-view classification is visible in Fig. 13 where some areas of the wound, which were misclassified in one of the two views, are well classified after the merging step.
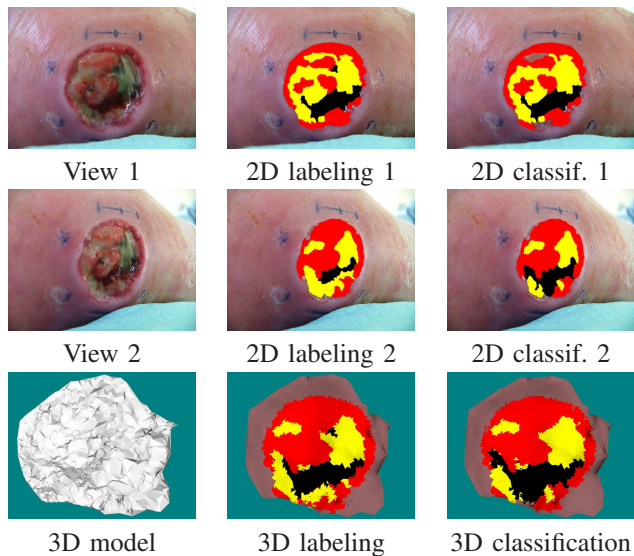


Fig. 13. Robustness of multi-view classification.

These tests show that the merging of 2D classifications enables more accurate tissue classification.

### C. Wound metric assessment

Several computations are possible on the labeled 3D model which is simply a mesh of elementary triangles.

Firstly, real tissue surfaces can be computed as the classification results are mapped onto the mesh surface of the 3D wound model. This is a substantial improvement as 2D assessment suffers from severe perspective distortions, preventing the accurate computation of tissue proportions.

Secondly, the multi-view classification enables the wound zone to be automatically isolated from the healthy skin, in order to extract the 3D wound model from the 3D body model captured (see (e) in Fig. 14). Formerly, it was impossible to extract exactly the wound from the 3D model because only a geometrical model and not a labeled one was available. The wound zone was roughly outlined using the mouse to specify the region in which points of interest where to be detected as input for the 3D reconstruction process. Obviously, not many of the points lay on the wound border, preventing an accurate contouring. Now, we need only to consider the triangles labeled as wound tissues and neighboring triangles labeled as healthy skin tissues. Each common line of the triangles is included in the wound outline. These lines constitute a closed boundary, making it possible to compute the 3D wound perimeter. Note that in the case where a triangle edge crosses the wound border, the classification results will display healthy and injured tissues and so the triangle will be split to fit to the wound border.

Finally, the inner volume of a wound, an interesting clinical index for the assessment of the healing process, can also be computed. It is extracted from the labeled 3D model by closing the wound surface with a reference plane. Of course, in the case where the wound is placed on a curved part of the body, typically on the heel, this plane should be replaced by a surface closer to the anatomical shape. Formerly, the reference plane was estimated from 3D points detected in a strip of healthy skin around the tracing [45] (see (d) in Fig. 14). This plane is now obtained by a least square minimization on the real 3D wound outline. The triangulation-based volume calculation consists simply in summing the elementary volumes under each prism formed by facets and their orthogonal projections on the reference plane (see (f) in Fig. 14). Another important clinical index is derived by closing the wound volume, namely the maximum depth over the wound.
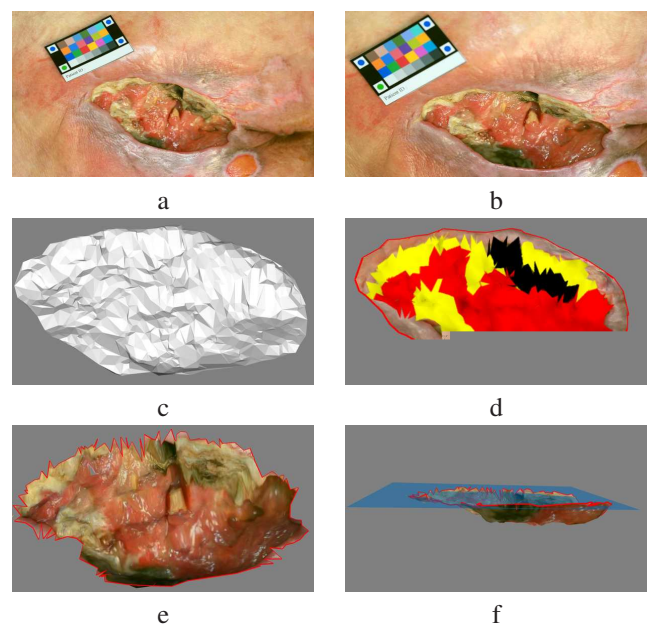


Fig. 14. Computing of real measurements by multi-view approach. (a-b) two views of the same wound (c) the 3D model calculated from these views (d) multi-view classification result (e) the detected zone of the wound isolated automatically from the healthy skin (f) the inner volume of the wound determined by closing the wound surface with a plane.

We can see in Table XI for the example given in Fig. 14 that the different wound measurements are significantly modified when the exact wound border is considered. 3D surface and volumes are good indices for therapeutic follow-up and clearly, tedious processes such as molding or serum injection will be advantageously avoided using 3D wound imaging.

| Index | No detection | Detection | Var. (%) |
|---|---|---|---|
| 3D surface ($cm^2$) | 45.09 | 29.28 | **-35.0** |
| Volume ($cm^3$) | 13.84 | 9.34 | **-32.5** |
| Max. depth (cm) | 1.49 | 1.55 | **+4.0** |

TABLE XI
COMPARISON OF SEVERAL WOUND INDICES WITHOUT AND WITH WOUND BORDER DETECTION.

## VI. Conclusion

Tissue classification is an essential part of wound assessment. By combining color and texture descriptors as an input vector of a SVM classifier, wound tissues have been classified from a single view into granulation, slough and necrotic tissues, competing with experts who had not seen the patients. However, a lack of repeatability is observed, due to the variations in the lighting conditions and the camera pose from one visit to the next. By combining 3D wound surface measurements with tissue classification in a single innovative and user-friendly tool, enhanced wound healing assessment has been achieved and as only a simple hand held digital camera is necessary, its widespread use by clinical staff will be very easy. The power of this strategy relies on the 3D wound model onto which tissue labels are mapped and classification results are merged. Experimental tests demonstrate that enhanced repeatability and robustness are obtained for tissue classification and that metric assessment is achieved through real area measurements, wound outline extraction and inner volume computation. In telemedicine environments, such a standardized and reproducible assessment of wounds using a digital camera is an essential requirement. A stronger agreement with the medical reference is obtained by multi-view classification results compared to single-view results, as observed on the overlap scores (73.8 against 69.9) and the Kappa coefficients (0.79 against 0.77). Finally, as demonstrated by the moderate inter-expert overlap scores, it should be noted that images alone cannot provide a complete understanding of wound healing and that the professional experience and patient knowledge of the clinician will remain essential for therapeutic decisions.

The extension of the current wound database is currently under way for enhanced tissue characterization and clinical staff education. In practical terms, a web site is dedicated to wound image uploading by clinical staff. We also intend to improve these results by matching regions from more than two views and by testing color descriptors that are invariant to viewpoint and lighting conditions. As many wounds are more complex and need further differentiation beyond the red/yellow/black scale, we are now investigating multispectral imaging capabilities.

The 3D model could also help classification, as granulation tissue is often bulb-shaped. Finally, the application of this complete wound assessment tool is currently in progress through clinical practise involving several care centers at the national level.

## Acknowledgments

## References

[1] A. J. Alberg, J. W. Park, B. W. Hager, M. V. Brock, and M. Diener-West, "The use of "overall accuracy" to evaluate the validity of screening or diagnostic tests," *Journal of General Internal Medicine*, vol. 19, pp. 460–465, 2004.

[2] A. Albouy, E. Koenig, S. Treuillet, and Y.Lucas:, "Accurate 3d structure measurements from two uncalibrated views," in *ACIVS*, 2006, pp. 1111–1121.

[3] B. Albouy, Y. Lucas, and S. Treuillet, "3d modeling from uncalibrated color images for a complete wound assessment tool." *Conf Proc IEEE Eng Med Biol Soc*, vol. 2007, pp. 3323–3326, 2007.

[4] B. Albouy, S. Treuillet, and Y. Lucas, "Robust semi-dense matching across uncalibrated and widely separated views," in *MVA 2007 IAPR Conference on machine vision applications University of Tokyo Japan*, University of Tokyo, May 2007, pp. 21–24.

[5] D. Beeckman, L. Schoonhoven, J. Fletcher, K. Furtado, L. Gunningberg, H. Heyman, C. Lindholm, L. Paquay, J. Verdu, and T. Defloor, "Epuap classification system for pressure ulcers: European reliability study," *Journal of Advanced Nursing*, vol. 60, no. 6, pp. 682–691, December 2007.

[6] B. Belem, "Non-invasive wound assessment by image analysis," Ph.D. dissertation, University of Glamorgan, Wales - UK, 2004.

[7] P. Berris, "Acquisition of skin wound images and measurement of wound healing rate and status using color image processing," Ph.D. dissertation, The University of Reading, United Kingdom, 2000.

[8] S. Bottou, F. Van den Heuvel, A. Cohen, and R. Scholtens, "Photogrammetric wound measurement with a three-camera vision system," in *Int. Archives of Photogrammetry and Remote Sensing*, vol. 33, Amsterdam, 1994, pp. 84–91.

[9] M. Callieri, P. Cignoni, M. Coluccia, G. Gaggio, P. Pingi, M. Romanelli, and R. Scopigno, "Derma : monitoring the evolution of skin lesions with a 3d system," in *8th Int. Workshop on vision, modeling and visualization*, Munich, 2003, Nov 19-21 2003, pp. 167–174.

[10] M. Celebi, Y. Aslandogan, and P. Bergstresser, "Unsupervised border detection of skin lesion images," in *Proc. Int. Conf. on information technology: coding and computing ITCC05*, vol. 2, 2005, pp. 123–128.

[11] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing kernel parameters for support vector machines," *Machine Learning*, pp. 131–160, 2001.

[12] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[13] O. Cula, K. Dana, F. Murphy, and B. Rao, "Skin texture modeling," *Int. Journal of Computer Vision*, vol. 62, no. 1-2, pp. 97–119, 2005.

[14] T. Defloor and L. Schoonhoven, "Inter-rater reliability of the epuap pressure ulcer classification system using photographs," *Journal of Clinical Nursing*, vol. 13, pp. 952–959, 2004.

[15] Y. Deng and B. S. Manjunath., "Unsupervised segmentation of color-texture regions in images and video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23(8), pp. 800–810, August 2001.

[16] A. Fadzil, M. Hani, N. Eltegani, S. Hussein, A. Jamil, and P. Gill, "Assessment of ulcer wound size using 3d skin surface imaging," *Lecture Notes in Computer Science*, vol. 5887, pp. 243–253, 2009.

[17] P. Felzenszwalb and D. Huttenlocher., "Efficient graph-based image segmentation," *Int. Journal of Computer Vision*, vol. 59(2), pp. 167–181, September 2004.

[18] M. Galushka, H. Zheng, D. Patterson, and L. Bradley, "Case-based tissue classification for monitoring leg ulcer healing," in *Proc. 18th IEEE Symposium on computer-based medical systems*, 2005, pp. 353–358.

[19] Y. V. Haeghen, J. M. Naeyaert, I. Lemahieu, and W. Philips, "An imaging system with calibrated color image acquisition for use in dermatology." *IEEE Trans Med Imaging*, vol. 19, no. 7, pp. 722–730, Jul 2000.

[20] G. Hansen, E. Sparrow, J. Kokate, K. Leland, and P. Iaizzo, "Wound status evaluation using color image processing," *IEEE Transactions on Medical Imaging*, vol. 16, no. 1, pp. 78–86, Feb. 1997.

[21] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.

[22] C.-W. Hsu and C.-J. Lin., "A comparison of methods for multi-class support vector machines." *IEEE Trans. on Neural Networks*, vol. 13, pp. 415–425, 2002.

[23] B. Jones and P. Plassmann, "An instrument to measure the dimensions of skin wounds," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 5, pp. 464–470, 1995.

[24] C. D. Jones, P. Plassmann, R. F. Stevens, M. R. Pointer, and M. B. McCarthy, "Good practice guide to the use of mavis ii," Medical Imaging Research Unit Technical Report TR-07-06, University of Glamorgan, Tech. Rep., 2006.

[25] M. Kolesnik and A. Fexa, "Multi-dimensional color histograms for segmentation of wounds in images," *Lecture Notes in Computer Science*, vol. 3656, pp. 1014–1022, 2005.

[26] D. Kosmopoulos and F. Tzevelekou, "Automated pressure ulcer lesion diagnosis for telemedicine systems," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, no. 5, pp. 18–22, 2007.

[27] T. A. Krouskop, R. Baker, and M. S. Wilson, "A noncontact wound measurement system." *J Rehabil Res Dev*, vol. 39, no. 3, pp. 337–345, 2002.

[28] J. Landis and G. Koch, "The measurement of observer agreement for categorical data." *Biometrics*, vol. 33, pp. 159–174, 1977.

[29] X. Liu, W. Kim, R. Schmidt, B. Drerup, and J. Song, "wound measurement by curvature maps: a feasability study," *Physiol. Meas.*, vol. 27, no. 11, pp. 1107–1123, nov 2006.

[30] J. Lowery, J. Hamill, E. Wilkins, and E. Clements, "Technical overview of a web-based telemedecine system for wound assessment," *Advances in Skin and Wound Care*, vol. 15, no. 4, pp. 165–169, August 2002.

[31] D. Lubeley, K. Jostschulte, R. Kays, K. Biskup, and C. B., "3d wound measurement system for telemedical applications," in *39th annual congress of the german society for biomedical engineering*, vol. 50, owner = hazem.wannous, timestamp = 2008.10.09, 2001.

[32] A. Malian, F. Heuvel van den, and A. Azizi, "A robust photogrammetric system for wound measurement," in *International Archives of Photogrammetry and Remote Sensing*, vol. 34, Corfu, Greece, 2-6 September 2002, pp. 264 –269.

[33] M. A. Malian, A. Azizi, and F. Van Den Heuvel, "Medphos: A new photogrammetric system for medical measurement," *Int. Archives of Photogrammetry and Remote Sensing*, vol. XXXV, pp. 929–933, 2004.

[34] H. Mesa, L. Morente, and F. Veredas, "Tissue recognition for pressure ulcer evaluation," in *ECIFMBE Proceedings*, S.-V. Berlin, Ed., vol. 22, 2008, pp. 1524–1527.

[35] H. Oduncu, A. Hoppe, M. Clark, R. J. Williams, and K. G. Harding, "Analysis of skin wound images using digital color image processing: a preliminary communication." *Int. Journal on Low Extremity Wounds*, vol. 3, no. 3, pp. 151–156, Sep 2004.

[36] C. Ozturk, S. Dubin, M. Schafer, W.-Y. Shi, and M. Chou, "A new structured light method for 3-d wound measurement," in *Proc. IEEE 22th annual northeast bioengineering Conf.*, 1996, pp. 70–71.

[37] A. Perez, A. Gonzaga, and J. Alves, "Segmentation and analysis of leg ulcers color images," in *Proc. Int. Workshop on medical imaging and augmented reality*, 10–12 June 2001, pp. 262–266.

[38] P. Plassmann, "Measuring wounds - a guide to the use of wound measurement tetchniques," *Journal of Wound Care*, vol. 4, no. 6, pp. 269–272, 1995.

[39] P. Plassmann and T. Jones, "Improved active contour models with application to measurement of leg ulcers," *SPIE Journal of Electronic Imaging*, vol. 12, no. 2, pp. 317–326, April 2003.

[40] ——, "Mavis: a non-invasive instrument to measure area and volume of wounds." *Medical Engineering and Physics*, vol. 20, pp. 332–338, July, 1998.

[41] L. Priese and V. Rehrmann, "Fast and robust segmentation of natural color scenes," in *Proceedings of asian conf. on computer vision*, Hongkong, 1998, pp. 598–606.

[42] M. Romanelli, G. Gaggio, M. Coluccia, F. Rizzello, and A. Piaggesi, "Technological advances in wound bed measurements," *Wounds*, vol. 14, pp. 58–66, 2002.

[43] H. Sanada, "Reliability and validity of DESIGN, a tool that classifies pressure ulcer severity and monitors healing," *Journal of Wound Care*, vol. 13, no. 1, pp. 13–18, 2004.

[44] C. Serrano, A. Begona, R. Rangaraj, and L. M. Roa, "Segmentation-based lossless compression of burn wound images," *Journal of Electronic Imaging*, pp. 720–726, 2001.

[45] S. Treuillet, B. Albouy, and Y. Lucas, "Three-dimensional assessment of skin wounds using a standard digital camera," *IEEE Transactions on Medical Imaging*, vol. 28, pp. 752–762, May 2009.

[46] P. Unger, C. Fife, and D. Weir, "Capturing the essence of the wound evaluation," *Today's Wound Clinic*, pp. 38–42, Spring 2008.

[47] H. Wannous, Y. Lucas, and S. Treuillet, "Supervised tissue classification from color images for a complete wound assessment tool," in *29th Int. Conf. of IEEE engineering in medicine and biology society - EMBS'07*, August 2007.

[48] ——, "Efficient svm classifier based on color and texture region features for wound tissue images," in *SPIE Medical Imaging: Computer-aided Diagnosis*, S. D. Library, Ed., vol. 6915 69152T, 17-21 Feb. 2008.

[49] H. Wannous, Y. Lucas, S. Treuillet, and B. Albouy, "A complete 3d wound assessment tool for accurate tissue classification and measurement," in *ICIP IEEE int. conf. on image processing*, October 2008, pp. 2928–2931.

[50] H. Wannous, S. Treuillet, and Y. Lucas, "Robust tissue classification for reproducible wound assessment in telemedicine environments," *Journal of Electronic Imaging*, vol. 19, no. 2, 2010.

[51] H. Zheng, L. Bradley, D. Patterson, M. Galushka, and J. Winder, "New protocol for leg ulcer tissue classification from colour images," in *Proc. 26th annual int. Conf. of the IEEE engineering in medicine and biology society*, vol. 1, 2004, pp. 1389 – 1392.