Annals of the BMVA Vol. 2012, No. 1, pp 1–14 (2012)

Histogram-based Motion Segmentation and Characterisation of Focal Liver Lesions in CEUS

Spyridon Bakas[1], Katerina Chatzimichail[2], Andreas Hoppe[1], Vasileios Galariotis[2], Gordon Hunter[1], Dimitrios Makris[1]

1. Digital Imaging Research Centre,
Faculty of Science, Engineering and Computing, Kingston University
Penrhyn Road, Kingston upon Thames, Surrey, KT1 2EE, London, United Kingdom
2. Radiology & Imaging Research Center
University of Athens, Evgenidion Hospital
Papadiamantopoulou Street 20, T.K.115 28, Athens, Greece.
⟨s.bakas@kingston.ac.uk, katerina@hcsl.com,
a.hoppe@kingston.ac.uk, vgalariotis@hotmail.com,
g.hunter@kingston.ac.uk, d.makris@kingston.ac.uk⟩

Abstract

The aim of this work is to provide an informative evaluation tool to assist clinicians diagnose focal liver lesions (FLLs) in Contrast-Enhanced Ultrasound (CEUS). A methodology to track and characterise a single FLL is presented. We propose a histogram-based motion segmentation approach, in combination with Lowe's SIFT keypoints, to track the locations of the FLL's enhancement region and the healthy liver throughout a CEUS video sequence, whilst intensity appearance is dramatically changed due to the effect of contrast agents. Then, Generalised Procrustes Analysis estimates the FLL's mean shape in order to characterise it as benign, or malignant. Finally, the precision of the automatically segmented areas of the FLL and the parenchyma are quantitatively analysed, to evaluate the level of confidence of our algorithm's decision on the characterisation of the FLLs included in our clinical data.

1 Introduction

Hepatic disease has a continuously increasing impact on mortality rates during the last 40 years rising by almost 385% over that period [CancerResearchUK, 2011] and is currently the fifth largest cause of death in the UK [British.Liver.Trust, 2011]. The group of hepatic disease is described by a number of medical conditions affecting the liver, e.g. hepatitis, cirrhosis, hepatocellular carcinoma (HCC), etc. These conditions can be either relatively

harmless conditions (benignities), or progressively worsening diseases that potentially result in death (malignancies). HCC is a distinctive example of a malignant focal liver lesion (FLL) and is the fifth most common type of cancer [Llovet et al., 2003]. Focal liver lesions are solid or liquid-containing nodules, "foreign" to the liver's anatomy. Extreme interest attracts the ability to distinguish early any case of premature (small) malignant FLLs from benign FLLs, as they can be healed without performing any surgical operation. This would result to an improvement on patients' care and eventually patients' survival prospects.

Visualisation of candidate FLLs has been attempted by employing different diagnostic imaging techniques throughout the years. Ultrasonography (US) has limited sensitivity in the detection of very small masses [Harvey and Albrecht, 2001] and provides vague results due to low image resolution and low signal-to-noise ratio when compared to Computed Tomography (CT) [Jamis-Dow et al., 1996] and Magnetic Resonance Imaging (MRI) [Itai et al., 1985]. CT and MRI have been used to clarify the US findings [Lencioni et al., 2007] and confirm a lesion's existence, as clearer images are obtained through these techniques. However, the cost of either a CT, or an MRI examination is a major disadvantage [Sirli et al., 2010]. Other disadvantages might include the inconveniently cumbersome size of the equipment used, as well as the exposure to emitted electromagnetic radiation, in the case of CT. Recently, Contrast-Enhanced Ultrasound (CEUS) has gained acceptance in the detection and characterisation of very small FLLs [Wilson and Burns, Oct. 2010]. Both diagnostic performance and confidence of CEUS in the characterisation of FLLs are found to be improved in comparison with baseline conventional US [Quaia et al., 2004]. In addition, given a significant improvement in sensitivity and specificity, CEUS is also supported as the reliable replacement of CT and MRI in the characterisation of HCC [Bruix et al., 2005], as well as of any metastatic liver diseases [Dietrich et al., 2006].

There is a rising need to use CEUS and create an associated software tool for clinicians and radiologists, primarily for two reasons. Firstly, CEUS is effective in terms of diagnostic accuracy that exceeds 95% for the evaluation of malignant FLLs, according to studies by radiologists [Strobel et al., 2009]. In addition, CEUS equipment is relatively inexpensive and portable when compared to other techniques (e.g. CT, MRI) [Sirli et al., 2010].

The introduction of contrast-enhancing agents into the human body during a CEUS examination results in changes of appearance of the FLL and the surrounding areas (Section 2). Specifically, their apparent 2D size, shape and intensity dramatically change over the few minutes of the procedure (see Fig. 1). As a consequence, the tracking and the characterisation of an FLL poses a very challenging task. Moreover, the potential instability of the clinician's hand that holds the transducer, as well as the patient's irregular breathing and the motion of the inner human organs, are external factors that affect the location of the FLL within the image. Furthermore, acoustic shadows, US absorption and low signal-to-noise ratio inevitably degrade the quality of the acquired data.

This paper describes a methodology to track an FLL and the rest of the healthy liver (parenchyma) over the time of a CEUS examination and how such information may assist a clinician's diagnosis. More specifically, we propose a histogram-based segmentation technique to track the extend of an FLL's enhancement region in subsequent frames, which allows differentiation between the intensity values of the lesion and the parenchyma. When the contrast between the FLL and the parenchyma is low, histogram-based segmentation is not feasible and the point-based registration technique of Lowe's scale-invariant feature transform (SIFT) [Lowe, 2004] is used to detect, localise and describe salient points in subsequent frames. Then, by establishing correspondences between the frames, the lesion is

Annals of the BMVA Vol. 2012, No. 1, pp 1–14 (2012)

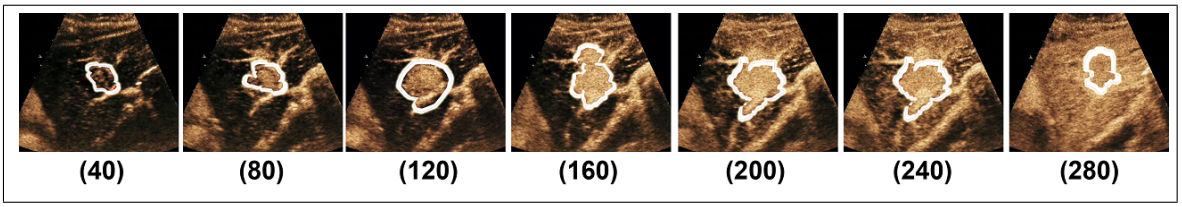


Figure 1: Focal liver lesion (FLL) tracking during the arterial phase video sequence, where the appearance of the liver and the lesion changes significantly. The third image (frame 120) is the frame where the FLL mask is initialised.

tracked over time, overcoming the dramatic appearance changes during a CEUS examination (Fig. 1). The lesion's silhouette descriptors, derived by tracking, are combined with the statistical analysis method of Generalised Procrustes Analysis (GPA) [Gower, 1975] to model the lesion's shape. Finally, size and shape information are used to localise the lesion during the last phase of the examination and characterise it, after taking into consideration pre-defined vascular signatures (VSs) [Rognin et al., 2009].

2 Contrast Enhanced Ultrasonography

CEUS is based on the intravenous injection of microbubble contrast agents, whose mechanism of action allows a confident diagnosis to be performed. Contrast agents, first introduced in the late nineties, offer a display enhancement in graylevel intensity, by maximising the contrast between an FLL and parenchyma. Agents oscillate or disrupt when exposed to a low or high-intensity ultrasound field respectively, allowing for an excellent depiction of the FLL vascularity and perfusion [Wilson and Burns, Oct. 2010]. Although a reliable diagnosis can be performed within ten minutes [Claudon et al., 2008], which is the maximum total duration of a CEUS examination [Albrecht et al., 2004], few radiologists have been trained to perform this modality and interpret its visual cues for diagnosis of FLLs.

A CEUS examination comprises three phases, whose durations vary, mainly depending on the physiopathology of the patient's liver and heart. After the intravenous injection of the contrast medium, the change of graylevel intensity with respect to the time elapsed represents the vitality of an FLL, by screening the inflow and outflow of the medium itself through the blood. The three phases of the examination are named considering the fact that the liver is supplied with blood initially from the hepatic artery and then from the portal vein. The first phase of CEUS, (arterial or enrichment phase) is characterised by the enrichment (increment in brightness) of the liver area due to the contrast agents. During the second phase (portal venous phase) the concentration and flow of contrast agents are stabilised, therefore there is no relative change in brightness. Finally, the third phase (late, acidic or parenchymatic phase) is characterised by the outflow of the agents from the lesion and the parenchyma (decrement in brightness).

There are two main VSs of FLLs; the hyper-enhancement and the hypo-enhancement. VSs describe the way that an FLL dynamically behaves in comparison to the parenchyma. During the first phase, the hyper-enhanced FLLs are enriched prior to the parenchyma, while in the hypo-enhanced category the parenchyma is enriched first. Lesions whose VS is hyper-enhancement throughout the three phases, are called unipolar hyper-enhanced. In contrast,

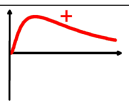
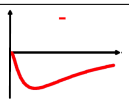
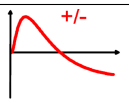
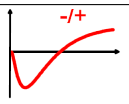
ID	VASCULAR SIGNATURE (VS)	SIGNAL	PHASE 1	PHASE 3	TYPICALLY FOUND IN
A	**Unipolar hyper-enhanced** (hyper-enhancement)		Lesion brighter than Parenchyma	Lesion brighter than Parenchyma	**Benign FLLs** (e.g. Adenoma, Abscess, Haemangioma, Focal Nodular Hyperplasia)
B	**Unipolar hypo-enhanced** (hypo-enhancement)		Lesion darker than Parenchyma	Lesion darker than Parenchyma	**Benign FLLs** (e.g. Fibrous tumour, Regenerating or dysplastic nodules, Simple cysts)
C	**Bipolar hyper-enhanced** (hyper-enhancement followed by hypo-enhancement)		Lesion brighter than Parenchyma	Lesion darker than Parenchyma	**Malignant FLLs** (e.g.Hepatocellular Carcinoma, Metastasis, Cholangiocarcinoma)
D	**Bipolar hypo-enhanced** (hypo-enhancement followed by hyper-enhancement)		Lesion darker than Parenchyma	Lesion brighter than Parenchyma	**Benign FLLs**

Table 1: Vascular signatures of FLLs relative to the parenchyma. [Adapted from [Rognin et al., 2009]]

bipolar hyper-enhanced are those where the contrast agents flow out of the lesions prior to the parenchyma during the third phase. If the VS of an FLL is hyper-enhanced, it is mainly in the third phase where its type (benign or malignant) is identified [Goertz et al., 2010], whereas if the VS is hypo-enhanced, it is in the first phase that the lesion's type is identified as benign [Rognin et al., 2009]. Therefore, the differentiation of a benign from a malignant lesion can be achieved by considering the alterations of the brightness intensity between the regions of the lesion and the parenchyma. Different VSs are linked to different medical conditions. Therefore, assigning an FLL to one of the four categories in Table 1 is of particular medical importance and assists the radiologist make a reliable diagnosis of an FLL of typical behaviour.

3 Previous studies

[Wilson and Burns, 2006] proposed a diagnostic process to radiologists to effectively characterise an FLL of typical behaviour, while using the CEUS modality. More specifically they described how characterisation (benign or malignant) may be based on observations of the FLL's average intensity enhancement in relation to the parenchyma's, during the third phase. Then, further distinction of the FLLs is allowed based on observations during the first phase enhancement patterns.

[Noble, 2010] discussed the trend of convergence of tissue characterisation and image segmentation in B-mode US for distinguishing between healthy and diseased tissues. Similarly, in our work, we propose an image segmentation method for characterising tissues in CEUS.

[Huang-Wei et al., 2006] have analysed video sequences of the arterial phase with an approximate duration of 7-12 seconds using the QontraXt software (AMID, Italy). Information was obtained after the application of a background subtraction technique [Piccardi, 2004], and then a colour mapping is used for improving visualisation. However, background subtraction may work only if the transducer is considered static and the patient's breathing is neglected. Also, the intensity values of the background (parenchyma) change continuously,

Annals of the BMVA Vol. 2012, No. 1, pp 1–14 (2012)

which further constrains the applicability of background subtraction. Last but not least, only Focal Nodular Hyperplasia lesions were analysed, which are unipolar hyper-enhanced lesions.

[Shiraishi et al., 2008] developed a computer-aided diagnostic scheme for the classification of three very specific types of FLLs with hyper-enhanced VS (i.e. Metastasis, Haemangioma and Hepatocellular Carcinoma). A cascade of six independent Artificial Neural Networks was proposed for the classification of the FLLs, based on manually segmented lesion regions in all frames of the acquired data.

Our approach attempts to track the FLL region through a CEUS video sequence to allow processing of cases despite transducer motion or patient's irregular breathing patterns. All recordings used in the proposed method are obtained without any standard criteria being instructed to the radiologist beforehand. Manual input from the user is only needed to initialise our tracking algorithm in phase 1. In addition to our previous work [Bakas et al., 2011], the proposed method is able to characterise FLLs of both hypo- and hyper-enhanced VS. Furthermore, the parenchyma region is also tracked, in order to provide more concise intensity information and result in a more accurate characterisation.

4 Methodology

The proposed method deals with FLLs of all different VSs and attempts to differentiate benign from malignant cases. Firstly, the video sequence of the arterial phase is processed, obtaining silhouettes of the FLL and the parenchyma for each frame of the sequence. A histogram-based motion segmentation technique is introduced, in combination with SIFT [Lowe, 2004], to track these silhouettes over the frames of the sequence. Subsequently, GPA [Gower, 1975] is employed to estimate the shapes of the lesion and parenchyma. The exhaustive search method of sliding window is then used to localise the lesion in a static image of the late phase. For every frame, the difference of the spatial average intensities of the lesion and the parenchyma determine the values of the signal of the lesion's VS. Identifying the type of the VS (Table 1) assists the characterisation of the lesion as benign or malignant.

The video sequence of the arterial phase is processed, to obtain the area and shape descriptors of the lesion and the parenchyma. Initially, the conical area viewed by CEUS is selected and, after removing any artefacts, is set as the mask of our workspace (US mask). These artefacts refer to regions whose brightness intensity has been enhanced prior to the examination, by the radiologist adjusting the controls of the US scanner, and therefore are of no interest. The FLL and the parenchyma masks are initialised on a single frame (t_0), ideally the one with the maximum contrast between lesion and parenchyma (Fig. 1). The initialisation process is performed manually by the radiologist, as prior medical knowledge is considered advantageous in correctly spotting a tumour, based on the perceived local intensity and shape information. Then, a Gaussian filter is applied to the region of the US mask to reduce the speckle noise in the image e.g. large intensity variance within neighbouring pixels.

The FLL and the parenchyma masks are tracked backwards and forwards in time from frame t_0 (Fig. 1). Firstly, a histogram-based segmentation is applied in subsequent frames, as described hereby. The intensity values of the FLL in frame t are represented by histogram $H_{F,t}$. Then intensity values in next frame $I_{(t+1)}$ are weighted according to $H_{F,t}$ to highlight

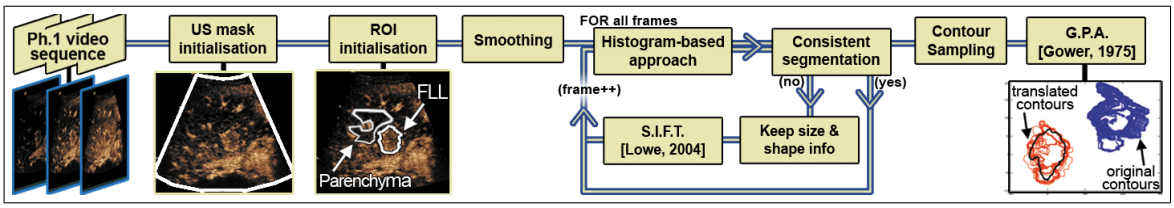


Figure 2: Diagram showing the procedure followed during Phase 1.

the FLL region in frame $t+1$.

$$p'_{x,y,(t+1)} = H_{F,t}(p_{x,y,(t+1)}) \tag{1}$$

Note that histogram $H_{F,t}$ and intensity values $p'_{x,y,(t+1)}$, $p_{x,y,(t+1)}$ are normalised between $[0,1]$. According to equation 1, parenchyma pixels are expected to be of intensity zero, while FLL pixels will have values proportional to the frequency of their intensities in the $H_{F,t}$ histogram. As a consequence, a simple threshold Z_{t+1} specified as the average intensity of all N pixels within the US mask (Equation 2) is in most cases sufficient to segment the FLL from the parenchyma.

$$Z_{t+1} = \frac{\sum_{i=1}^{N} p'_{i,(t+1)}}{N} \tag{2}$$

The changes of the intensity histograms of the lesion and the parenchyma are considered to be insignificant over consecutive frames. The thresholded image is cleaned using morphological opening with a disk-shaped structuring element, in order to remove very small regions, protrusions from objects and thin connections between objects, all of which are likely to be artefacts due to noise. Subsequently, the resulting image is segmented using the connected components algorithm. If this segmentation result is consistent with the FLL region of frame t, then it is set as the FLL region in frame $t+1$. The consistency of the segmentation result is verified by the following two conditions [Bakas et al., 2011]:

a Relative size difference: $\frac{|S_{F,t}-S_{F,(t+1)}|}{S_{F,t}} < D_s$, where $S_{F,t}$ is the size of the FLL at frame t,

b Displacement of centre of gravity: $\sqrt{(C_{x,t}-C_{x,t+1})^2+(C_{y,t}-C_{y,t+1})^2} < D_d$, where $C_{x,t}$ and $C_{y,t}$ are the x and y coordinates of the FLL's centre of gravity at frame t.

However, the resulting FLL region in frame $t+1$ may not be consistent with frame t, because of low contrast between the FLL and the parenchyma, and/or the movement of the transducer. In such cases, SIFT is used to capture the changes in the location of the FLL's centre of gravity, by localising matching sets of salient points Q_t and Q_{t+1} in two consecutive frames. Q_t is a $2 \times K_t$ matrix, comprising K_t 2D salient points $q_{\kappa,t}$ at time t, where $\kappa \in [1, K_t]$. Specifically, each $q_{\kappa,t}$ is assigned a statistical descriptor ($V(q_{\kappa,t})$) of 128 dimensions, in order to characterise it. Then the best correspondences of the detected keypoints between the two frames are found by minimising the nearest neighbour distance ratio (NNDR) between the descriptors of the two frames ($V(q_{\kappa,t})$, $V(q_{\kappa,t+1})$) and selecting the nearest point in the descriptor space. Specifically, for every point $q_{\kappa,t}$, another point $q_{\lambda,t+1}$, where $\lambda \in [1, K_{t+1}]$, is found that fulfills the following equation:

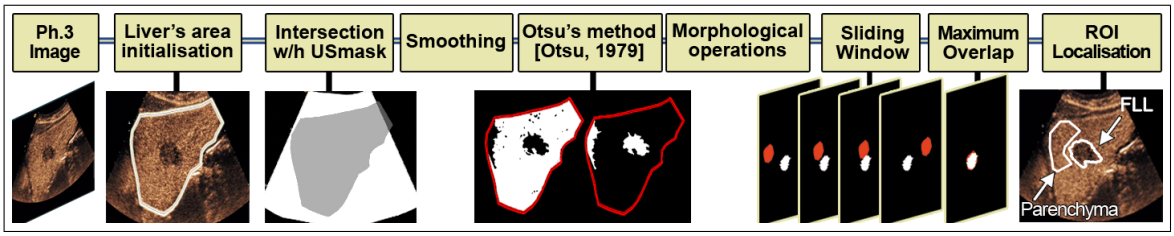


Figure 3: Diagram showing the procedure followed during Phase 3.

$$(q_{\kappa,t}, q_{\lambda,t+1}) = \underset{\substack{q_{\kappa,t}\in Q_t \\ q_{\lambda,t+1}\in Q_{t+1}}}{\operatorname{argmin}} [\arccos(V^T(q_{\kappa,t}) \bullet V(q_{\lambda,t+1}))] \tag{3}$$

Similarly, we find a second point $q_{\lambda',t+1}$, where $\lambda' \in [1, K_{t+1}]$, such that:

$$(q_{\kappa,t}, q_{\lambda',t+1}) = \underset{\substack{q_{\kappa,t}\in Q_t \\ q_{\lambda',t+1}\in Q_{t+1}, \lambda'\neq\lambda}}{\operatorname{argmin}} [\arccos(V^T(q_{\kappa,t}) \bullet V(q_{\lambda',t+1}))] \tag{4}$$

We accept that there is a correspondence between points $q_{\kappa,t}$ and $q_{\lambda,t+1}$ in the two frames t and $t+1$, if:

$$\frac{\arccos(V^T(q_{\kappa,t}) \bullet V(q_{\lambda,t+1}))}{\arccos(V^T(q_{\kappa,t}) \bullet V(q_{\lambda',t+1}))} < Z \tag{5}$$

where $Z \in [0,1]$ is a threshold value. If a point $q_{\lambda,t+1}$ is the best match for more than one point $q_{\kappa,t}$, then these correspondences of $q_{\lambda,t+1}$ are rejected. Consequently, we only keep reliable correspondences.

After finding these correspondences between the two frames, the shape and size of the FLL region in frame $t+1$ are considered to be the same as in frame t and its $C_{x,y,t}$ is translated by the average displacement of the SIFT keypoints, allowing localisation of the FLL. When the contrast between the FLL and the parenchyma is low, the histogram-based motion segmentation becomes unreliable (Fig. 1 - frame 280), while SIFT performs satisfactorily, as its keypoints are invariant to illumination changes [Lowe, 2004]. The parenchyma mask is translated by the same average displacement in each frame throughout the whole sequence and without considering any changes to its initialised shape. When tracking backwards in time, all $(t+1)$ notations in the above equations should be substituted by $(t-1)$.

The silhouettes of the lesion are obtained after applying the above steps to all frames of the arterial phase video sequence. The contours of the derived silhouettes are extracted and sampled to a set of marked points, i.e. a set of suitable "landmarks", in order to reduce the computational complexity, whilst representing satisfactorily the shape of the lesion. A fixed number of points is selected for all contours, to allow the application of the GPA. The mean shape of the FLL during the arterial phase is obtained by GPA, which optimally aligns the set of all contours in a single reference orientation. Such a shape may be more representative than the shape defined manually by the operator in a single frame t_0 as it takes into account evidence from multiple frames.

The mean shape (as estimated by GPA) is then used to localise the FLL in the static image of the late phase and assist the differentiation between benign or malignant case. Firstly, the liver area is manually specified on the image (Fig. 3) and its intersection with the US

D_s	10%			20%			30%			40%		
D_d	$5px$	$10px$	$15px$	$5px$	$10px$	$15px$	$5px$	$10px$	$15px$	$5px$	$10px$	$15px$
No Smoothing Application												
OL_F	70.0	69.5	69.6	65.9	65.9	65.9	62.5	61.5	62.2	61.7	61.7	61.5
OL_P	91.7											
Gaussian 5×5 filter ($\sigma = 5.0$)												
OL_F	75.9	75.9	75.0	75.5	75.5	74.6	74.3	73.9	73.9	73.7	73.7	73.7
OL_P	91.6											
Optimized Bayesian Non-Local Means Filter [Coupe' et al., 2009]												
OL_F	68.3	67.4	64.6	64.5	59.6	57.0	63.5	63.2	63.2	54.8	54.2	53.4
OL_P	92.0											

Table 2: Quantitative analysis of the automatically segmented regions with respect to their ground truth, after application of different smoothing techniques and different values of the consistency thresholds (D_d, D_s).

mask defines the ROI. The ROI is filtered with a Gaussian kernel and then Otsu's method [Otsu, 1979] is used to automatically select a threshold and binarise the ROI into foreground and background. Subsequently, the morphological operations of opening and closing are applied to the foreground and background areas respectively. The ROI is segmented using the connected components algorithm and small areas are removed. The FLL area is finally localised by maximising the intersection between the segmented areas (Fig. 3) and the mean shape, as estimated by GPA in the arterial phase, allowing translation and rotation of the latter through an exhaustive search (e.g. sliding window) (Fig. 3). Finally, the translation difference between C_{x,y,t_0} and C_{x,y,t_ψ} is calculated and applied to the parenchyma mask from frame t_0, in order to estimate its location in the late phase.

The differentiation between benign and malignant lesions is based on the signal of the lesion's VS ($F(t)$), calculated as:

$$F(t) = \frac{\sum_{i=1}^{M_t} p_{x_i,y_i,t}}{M_t} - \frac{\sum_{j=1}^{L_t} p_{x_j,y_j,t}}{L_t} \tag{6}$$

where M_t, L_t are the pixels within the FLL and parenchyma regions, respectively, for every frame t.

5 Experiments & Results

The equipment used for data acquisition included a Siemens ACUSON Sequoia C512 system, in combination with low-frequency 6C2 convex Transducer (2-6 MHz) capturing images at 25 frames-per-second. The second-generation contrast medium used was sulphur hexafluoride gas (SonoVue from Bracco Diagnostics). Acquisition parameters are unknown, as they were set by the radiologist separately for each patient. The captured data were exported as videos and images with resolution 768 x 576 pixels, with no compression applied.

The visual cues used are real clinical data of seventeen patients under similar physical conditions. Each case is described by a short video sequence of the arterial phase of 84-419 frames and one static image of the late phase, where the only assumption made was that only one lesion exists within the liver of each patient, with diameter between 3 and 6 cm. The "gold standard" is the clinician's decision on the characterisation of each lesion's type,

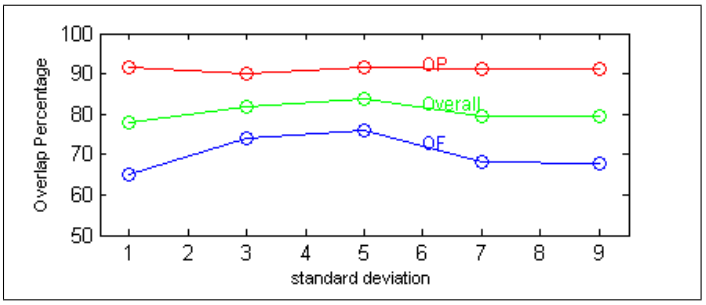


Figure 4: Quantitative analysis of the different Gaussian kernels' standard deviation (σ) applied for smoothing.

which has been based on pathology. Motion artefacts, acoustic shadows and US absorption inevitably degraded the quality of our imagery. Moreover, transducer movement and the patient's irregular breathing patterns also affected the quality of the recordings ("wiggly" recording). Specifically, in 7/17 cases (41%), the FLL was not continuously observed during the whole video sequence, because of either out-of-plane lesion movement or its dispersion in depth.

Results are quantitatively analysed by comparing the manually annotated ground truth (GT) of the FLL (G_{F_t}) and the parenchyma (G_{P_t}) with the automatically segmented areas of the FLL (A_{F_t}) and the parenchyma (A_{P_t}) throughout the sequence. The spatial overlap metric for the FLL regions (O_F) is computed by the formula $O_F = \frac{1}{T}\sum_{t=1}^{T} o_{F_t}$, where $o_{F_t} = \frac{|G_{F_t} \cap A_{F_t}|}{|A_{F_t}|}$ and T is the number of processed frames. This (o_{F_t}) obtains information only from pixels $p_{x,y,t} \in G_{F_t} \cap A_{F_t}$ and penalises only pixels misclassified as FLL. Similarly, the spatial overlap metric for the parenchyma regions (O_P) is computed as $O_P = \frac{1}{T}\sum_{t=1}^{T} o_{P_t}$, where $o_{P_t} = \frac{|G_{P_t} \cap A_{P_t}|}{|A_{P_t}|}$. This allows the operator to initialise a region ($A_{P_{t_0}}$), smaller than G_{P_t}, for tracking at frame t_0, as correct intensity information will still be captured within G_{P_t}, considering the spatial uniformity of the enrichment of G_{P_t}.

It is widely accepted that images acquired by the US modality are inherently very noisy including a lot of speckle noise [Coupe' et al., 2009]. Hence, it is difficult to interpret US images and successfully detect features, as neither edges are strongly defined, nor do areas that should be homogeneous appear so. On the other side, since speckle noise is associated with the tissue's response to US, its presence may actually assist the tissue tracking, due to the apparent texture that it causes. Table 2 presents the results of different smoothing options (no smoothing, Bayesian non-local means-based speckle filter [Coupe' et al., 2009], Gaussian filter) for a variety of parameters. The filter of [Coupe' et al., 2009] leads to results even worse than those obtained by omitting any smoothing, as SIFT fails to find any local features to track after the removal of the speckle noise in that case. The Gaussian filter that is adopted in our methodology not only smoothes out any Gaussian noise, but also preserves the eventually useful speckle noise and leads to higher tracking accuracy.

Subsequently, we tried Gaussian kernels of size 5-by-5 pixels, with different standard deviations ($\sigma \in [1,3,5,7,9]$). Results of the application of these filters to our clinical data, depict the kernel with standard deviation $\sigma = 5.0$ as the one giving the maximum overall

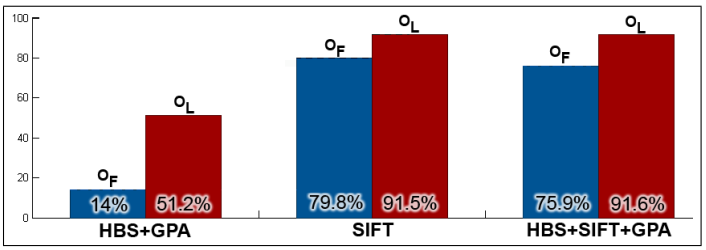


Figure 5: Comparative results of the different approaches applied during phase 1.

Case studies	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	%
Gold Standard	M	B	M	M	M	B	M	M	B	M	M	B	B	B	B	B	B	
SIFT	M	B	B	B	B	B	M	M	B	B	M	B	B	B	B	B	M	70.6
HBS+GPA	M	M	B	B	M	B	M	M	B	M	M	B	-	-	-	-	B	58.8
HBS+SIFT+GPA (Our method)	M	B	M	M	M	B	M	M	B	M	M	M	M	B	B	B	M	82.4

Table 3: Results on the characterised type of FLL, according to the different methods applied. The last column (%) refers to the successful characterisation percentage. (B::Benign, M::Malignant)

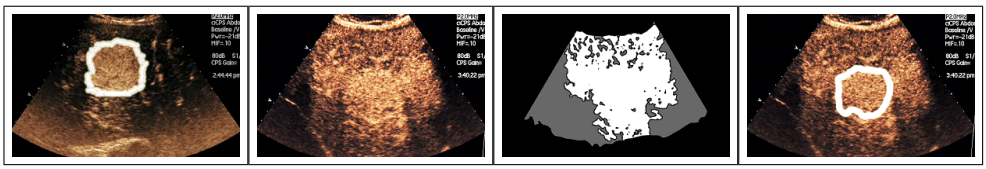
overlap percentage (83.8%) (Fig. 4). Fine tuning of the standard deviation may marginally improve the results.

The effect of the consistency thresholds (D_d and D_s) is also shown in Table 2. The combination of $D_d = 5$ pixels and $D_s = 10\%$ seems to optimise the tracking performance, independently of the smoothing method applied. However, relatively small variations appear in the performance of O_F and O_P for different values of the consistency thresholds.

The full version of the proposed methodology, as described in Section 4, obtained results for average spatial overlaps, O_F and O_P, equal to 75.9% and 91.6% respectively. The proposed method characterised successfully 14/17 cases (success rate 82.4%) when compared to the gold standard, as shown in Table 3. To highlight the importance of the different key elements of our method (Histogram-based motion segmentation, SIFT, GPA), we compare it with two simplified versions.

The first simplified version uses the histogram-based segmentation method for tracking the FLL during the 1st phase and then GPA for estimating the FLL's mean shape. This version adapts to changes of the FLL's shape, as segmentation is intensity-based. However, Fig. 5 shows that the intensity-based tracking has worse performance than the SIFT-based versions, as it struggles in low contrast frames. Similarly, the success rate of the FLL's characterisation is worse (Table 3) and in some cases impossible (cases 13-16), as the method failed to capture a reasonable shape for the FLL's enhancement region.

Tracking in the second simplified version is entirely based on SIFT. As explained in Section 4, SIFT estimates the average translation of the FLL's centre of gravity, without adjusting the FLL's shape. Therefore, as the initialised shape at frame t_0 is fixed over time, GPA is redundant for estimating the FLL's mean shape. Fig. 5 shows that this simplified version has marginally better tracking accuracy than the proposed methodology. However, the manually segmented FLL shape may not be so accurate as the one that is estimated by the histogram-based segmentation and GPA. Such a difference may lead to worse localisation



(a) (b) (c) (d)

Figure 6: Challenging case study depicting a malignant FLL within a cirrhotic liver. a) FLL initialisation (frame t_0). b) Late phase image. c) Expected FLL localisation. White area depicts the FLL and gray area the parenchyma. d) The proposed method precisely localising the FLL.

of the FLL in the late phase and therefore to worse characterisation performance, as seen in Table 3.

Therefore, the combination of histogram-based segmentation and SIFT manages to accurately track the location, shape and size of the FLL's enhancement region in the image plane throughout the video sequence of the first phase, despite the irregular breathing patterns of the patient and/or the irregular motion of the transducer. In addition, the estimation of the average FLL's shape by GPA improves the characterisation of the FLL when the third phase is considered.

The proposed method has failed in characterising 3/17 examined cases. These comprise cases 12, 13 and 17 from Table 3. Even though the VS of case 12 was correctly categorised as bipolar hyper-enhanced, the lesion was wrongly characterised as malignant due to its behaviour being atypical. Case 13's VS was categorised as bipolar hyper-enhanced due to increased non-rigid motion in the arterial phase video sequence related to the motion of the transducer and/or breathing, resulting in tracking failure. Finally, case 17 was not correctly characterised due to a very dark video sequence that caused failure of the proposed method.

The most challenging of the case studies examined included a malignant FLL within a cirrhotic liver (Fig. 6 and Table 3-case study 8). It was expected that the acquisition of the lesion's shape model and size information could be obtained during the arterial phase without any conflicts. However, this particular case was described as problematic by the radiologists, since the cirrhosis would make the localisation of the lesion and characterisation of its type during the late phase difficult. More specifically, as illustrated in Fig. 6(b), during the late phase both the parenchyma and the lesion appear brighter than the surrounding area. It is worth noting that even for an experienced specialist radiologist manually interpreting images of this nature, the task of identifying the lesion's type correctly, is not trivial in such cases, due to the same reason. Thus, it was expected that the method would assume almost all the liver to be one big FLL and the surrounding area to be the parenchyma. However, the proposed method first uses information of the lesion's size and shape from the arterial phase and then localises the lesion precisely in the late phase (Fig. 6(d)). By subsequently performing a spatial averaging of the brightness intensities of the lesion and the parenchyma, it was found that the VS of the lesion in this example is bipolar hyper-enhanced and the case was therefore correctly characterised as malignant.

6 Conclusions

This paper has introduced a methodology for tracking an FLL and the parenchyma over time, in order to obtain appropriate information needed to characterise it as either benign or malignant, based on its VS. Experimental results on real clinical case studies of different patients under identical physical conditions demonstrate 82.4% success rate on the characterisation of the currently available data, with the level of confidence that the obtained characterisation is truly representative of each FLL's type exceeding 83.5% overall (Fig. **??**).

There are three key elements used throughout the arterial phase video sequence that are essential for the success of the proposed method. First a segmentation process, based on histogram analysis of the region of the FLL and its consistency conditions [Bakas et al., 2011], is needed for estimating the changes in the size, shape and position of the lesion's enhancement region. Then, if these consistency conditions are not satisfied, SIFT is used for tracking a lesion, giving its silhouette at the same time in the frames where the segmentation process fails. Last, but not least, GPA is needed for reliable estimation of the lesion's mean shape. The combination of these elements provides adequate information to accurately localise the FLL during the late phase, categorise it to a VS and thus characterise its type as benign or malignant.

Our methodology requires the manual initialisation of the US area, the FLL and the parenchyma in a single frame t_0 of the arterial phase video sequence, as well as initialisation of the liver area on the late phase image. Automatic localisation of the above areas would allow a fully automatic method in place of the current semi-automated process with manual initialisation.

Additionally, in the future, we will investigate the different spatio-temporal patterns of the FLL enhancement during the arterial phase (e.g. centripetal, centrifugal, circular) [Wilson and Burns, Oct. 2010] and how to further support the diagnostic process.

References

T Albrecht, M Blomley, L Bolondi, M Claudon, J-M Correas, D Cosgrove, L Greiner, K Jager, N de Jong, E Leen, R Lencioni, D Lindsell, A Martegani, L Solbiati, L Thorelius, F Tranquart, H P Weskott, and T Whittingham. Guidelines for the use of contrast agents in ultrasound - january 2004. Ultraschall in der Medizin, 25(4):249–256, 2004.

S Bakas, K Chatzimichail, A Autret, A Hoppe, V Galariotis, and D Makris. Localisation and characterisation of focal liver lesions using contrast-enhanced ultrasonographic visual cues. In Proc. Medical Image Understanding and Analysis, 2011.

British.Liver.Trust. http://www.britishlivertrust.org.uk/home/looking-after-your-liver.aspx. Website, (Date of last access: 09 December 2011), 2011.

J Bruix, G Gores, and M Kojiro. Monothematic conference: Hepatocellular carcinoma - towards molecular profiling of clinical management. European Association for the Study of the Liver - American Association for the Study of Liver Disease - Japan Society of Hepatology, June 2005.

CancerResearchUK. http://info.cancerresearchuk.org/cancerstats/types/liver/mortality/. Website, (Date of last access: 09 December 2011), 2011.

M Claudon, D Cosgrove, T Albrecht, L Bolondi, M Bosio, F Calliada, J-M Correas, K Darge, C Dietrich, M D'Onofrio, D H Evans, C Filice, L Greiner, K Jager, N deJong, E Leen, R Lencioni, D Lindsell, A Martegani, S Meairs, C Nolse, F Piscaglia, P Ricci, G Seidel, B Skjoldbye, L Solbiati, L Thorelius, F Tranquart, H P Weskott, and T Whittingham. Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (ceus) - update 2008. Ultraschall in der Medizin, 29(1):28–44, 2008.

P Coupe', P Hellier, C Kervrann, and C Barillot. Nonlocal means-based speckle filtering for ultrasound images. IEEE Transactions on Image Processing, 18(10):2221–9, 2009.

C F Dietrich, W Kratzer, D Strobel, E Danse, R Fessl, A Bunk, U Vossas, K Hauenstein, W Koch, W Blank, M Oudkerk, D Hahn, and C Greis. Assessment of metastatic liver disease in patients with primary extrahepatic tumors by contrast-enhanced sonography versus ct and mri. World Journal of Gastroenterology, 12:1699–1705, 2006.

R S Goertz, T Bernatik, D Strobel, E G Hahn, and T Haendl. Software-based quantification of contrast-enhanced ultrasound in focal liver lesions - a feasibility study. European Journal of Radiology, 75:22–26, 2010.

J Gower. Generalized procrustes analysis. Psychometrika, 40:33–51, 1975.

C J Harvey and T Albrecht. Ultrasound of focal liver lesions. European Radiology, 11(9): 1578–1593, 2001.

C Huang-Wei, A Bleuzen, P Bourlier, J Roumy, A Bouakaz, L Pourcelot, and F Tranquart. Differential diagnosis of focal nodular hyperplasia with quantitative parametric analysis in contrast-enhanced sonography. Investigative Radiology, 41(3):363–368, 2006.

Y Itai, K Ohtomo, S Furui, T Yamauchi, M Minami, and N Yashiro. Non-invasive diagnosis of small cavernous hemangioma of the liver: Advantage of mri. American Journal of Roentgenology, 145(6):1195–1199, 1985.

C A Jamis-Dow, P L Choyke, S B Jennings, W M Linehan, K N Thakore, and M M Walther. Small $(< or = 3cm)$ renal masses: Detection with ct versus us and pathologic correlation. Radiology, 198(3):785–788, 1996.

R Lencioni, C D Pina, L Crocetti, E Bozzi, and D Cioni. Clinical management of focal liver lesions: the key role of real-time contrast-enhanced us. European Radiology Supplements, 17(6):F73–F79, 2007.

J M Llovet, A Burroughs, and J Bruix. Hepatocellular carcinoma. Lancet, 362:1907–1917, 2003.

D G Lowe. Distinctive image features from scale-invariant keypoints. Int. Journal of Computer Vision, 60(2):91–110, 2004.

J A Noble. Ultrasound image segmentation and tissue characterisation. Special Issue Paper Proc. Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 224 (2):307–316, 2010.

N Otsu. A threshold selection method from gray-level histograms. IEEE Trans. on Systems, Man, and Cybernetics, 9(1):62–66, 1979.

M Piccardi. Background subtraction techniques: a review. IEEE International Conference on Systems, Man and Cybernetics, 4:3099–3104, 2004.

E Quaia, F Calliada, M Bertolotto, S Rossi, L Garioni, L Rosa, and R Pozzi-Mucelli. Characterization of focal liver lesions with contrast-specific us modes and a sulfer hexafluoride-filled microbubble contrast agent: Diagnostic performance and confidence. Radiology, 232:420–430, 2004.

N G Rognin, L Mercier, P Frinking, M Arditi, G Perrenoud, A Anaye, and J Y Meuwly. Parametric imaging of dynamic vascular patterns of focal liver lesions in contrast-enhanced ultrasound. IEEE Int. Ultrasonics Symposium, pages 1282–1285, 2009.

J Shiraishi, K Sugimoto, F Moriyasu, N Kamiyama, and K Doi. Computer-aided diagnosis for the classification of focal liver lesions by use of contrast-enhanced ultrasonography. Medical Physics, 35(5):1734–1746, 2008.

R Sirli, I Sporea, A Martie, A Popescu, and M Danila. Contrast enhanced ultrasound in focal liver lesions - a cost efficiency study. Medical Ultrasonography, 12(4):280–285, 2010.

D Strobel, K Seitz, W Blank, A Schuler, C F Dietrich, A vonHerbay, M Friedrich-Rust, and T Bernatik. Tumor-specific vascularization pattern of liver metastasis, hepatocellular carcinoma, hemangioma and focal nodular hyperplasia in the differential diagnosis of 1349 liver lesions in contrast-enhanced ultrasound (ceus). Ultraschall in der Medizin, 30(4): 376–382, 2009.

S R Wilson and P N Burns. An algorithm for the diagnosis of focal liver masses using microbubble contrast-enhanced pulse-inversion sonography. American Journal of Roentgenology, 186:1401–1412, 2006.

S R Wilson and P N Burns. Microbubble-enhanced us in body imaging: What role? Radiology, 257(1):24–39, Oct. 2010.