

# Spatio-temporal Speckle Reduction in Ultrasound Sequences

Noura Azzabou<sup>1,\*</sup> and Nikos Paragios<sup>1,2</sup>

<sup>1</sup> Laboratoire MAS Ecole Centrale de Paris, France

<sup>2</sup> GALEN Group, INRIA Saclay, Ile-de-France, France

n.azzabou@institut-myologie.org, nikos.paragios@ecp.fr

**Abstract.** In this paper we will be concerned with speckle removal in ultrasound images. To this end, we introduce a new spatio-temporal denoising method based on a variational formulation. The regularization relies on a non parametric image model that describes the observed image structure and express inter-dependencies between pixels in space and time. Furthermore, we introduce a new data term adapted to the Rayleigh distribution of the speckle. The interaction between pixels is determined through the definition of new measure of similarity between them to better reflect image content. To compute this similarity measure, we take into consideration the spatial aspect as well as the temporal one. Experiments were carried on both synthetic and real data and the results show the potential of our method.

## 1 Introduction

Ultrasound imaging is a popular non invasive and low cost technique to observe the dynamical behavior of organs. The produced images inherit a multiplicative component that is the speckle. This signal may contain information useful to medical experts [1]. Nevertheless, for some applications like segmentation or registration the speckle affects the quality of images and complicates the diagnosis task. Therefore, speckle removal could be a tool to obtain better images while preserving details. In the literature, numerous methods for speckle suppression were proposed. Some of them transform the multiplicative noise problem into an additive one by considering the logarithm of the image and assume that the noise is Gaussian [2]. The limitation of such a method lies in the fact that the logarithm function reduces the contrast in the image and makes the task of denoising more complex. The anisotropic diffusion [3] relies on the gradient information which is not robust when dealing with high noise levels. We want to point out also that these methods, consider only fixed images and not sequences. The temporal aspect was addressed in [4] using temporal averaging. This technique is efficient tool of speckle removal but due to motion, fine details can be blurred unless an appropriate weight definition is done. Techniques that aim to perform signal correction through motion correction were also investigated [5]. Analysis

---

\* The first author is affiliated now to AIM -CEA NMR Laboratory Institut of Myology.

in the transform domain like wavelet was also considered toward speckle reduction [6,7], but the transform selection is crucial for the performance of these methods.

In this paper, we will address the speckle removal problem in ultrasound images. Inspired from total variation techniques, our approach entails a functional minimization process. Such a functional is composed of two terms : the regularization term that is able to encode complex image structure and the data term that encodes the speckle model. In addition to the image structure, the regularization term is designed to take into consideration the temporal aspect. In this context, the interactions between pixels are determined by weights that reflect the similarity in the temporal and the spatial domain with implicit constraints imposing motion consistency. As far as fidelity to data term is considered, its formulation is based on the speckle model which is assumed to be a Rayleigh distribution. The paper is organized as follows, section 2 is devoted to the presentation of the method, we present in section 3 the definition of weights (or interaction between pixels), next we will focus on the experimental validation of our approach in section 4, and finally we conclude the paper in section 5.

## 2 Problem Statement

Let us consider two image sequences  $U$  and  $I$  defined on  $T \times \Omega$  where  $T$  is the length of the sequence and  $\Omega$  refers to a frame spatial domain. These two sequences are related according to  $I = U * n$  with  $n$  being a multiplicative noise sequence that follows a Rayleigh distribution [1] defined as :  $r(n) = \frac{n}{\sigma_n^2} \exp(-\frac{n^2}{2\sigma_n^2})$ . In order to recover the noise free sequence  $U$  we have to minimize a cost function that insures smoothness, preserves boundaries, while imposing a fidelity constraint, to the observed sequence. To define the fidelity to data term, we will consider the link between the Maximum a Posteriori (MAP) estimation framework and the total variation formulation. The MAP estimator aims to minimize an energy composed of two terms : the image model that is the regularization term and the noise model that corresponds to the fidelity to data term. In the discrete case and for a frame  $U_t$  the observed noise likelihood is equivalent to the data term that can be expressed as

$$E_{data}(U_t) = -\log(P(I_t|U_t)) = -\log(P(n_t))$$

Under the assumption that (i) the noise observations are independent (ii) identically distributed, and (iii)  $U$  and  $I$  have non zero intensity for each pixel in the sequence, then we can write:

$$E_{data}(U_t) = -\log \left( \prod_{i=1}^{|\Omega|} r(n_t(i)) \right) = \sum_{i=1}^{|\Omega|} \left[ -\log \left( \frac{I_t(i)}{\sigma_n^2 U_t(i)} \right) + \frac{I_t^2(i)}{2\sigma_n^2 U_t^2(i)} \right]$$

Note that the assumption of independence between the noise observations is too simple for the speckle because it is correlated. However such an assumption yields

a simpler energy and a computationally efficient formulation. Now, if we consider the substitute the continuous formulation of the data term to the discrete one, we obtain

$$E_{data}(U_t) = \int_{\Omega} \left[ \log(U_t(\mathbf{x})) + \frac{I_t^2(\mathbf{x})}{2\sigma_n^2 U_t^2(\mathbf{x})} + Cte \right] d\mathbf{x} \tag{1}$$

Now, if we consider the function  $f(x) = \log(x) + \frac{1}{2\sigma_n^2} \frac{I_t^2(\mathbf{x})}{x^2}$ , it can be easily shown that the second derivative of  $f$  is positive for  $x \leq \frac{\sqrt{6}I_t(\mathbf{x})}{2\sigma_n^2}$ . This function is convex on the interval  $]0, 255]$  if  $\frac{\sqrt{6}I_t(\mathbf{x})}{2\sigma_n^2} > 255$  for any intensity value  $I_t(\mathbf{x})$ . In other words, if  $\left(\sigma_n^2 \leq \frac{\sqrt{6} \times \text{Inf}(I_t)}{2 \times 255}\right)$  then the designed function  $f$  is convex. In practice it is verified when  $(\sigma_n \leq 0.1)$  which is equivalent to small noise amount. Hence for small value of  $\sigma_n$ , the data term  $E_{data}$  is a convex function.

Regarding the regularization, we selected the model introduced in [8] that is a generalized linear interpolation model where each pixel is expressed as a weighted mean of the remaining pixels of the sequence. Such a model involves a data driven image model where the interaction between pixels is defined according to the similarity between them in order to preserve image content. For this application, we will consider the extension of this formulation to ultrasound sequences to account for the time dimension and to perform a temporal filtering as well as a spatial one. In this case  $E_{reg}$  is defined as

$$E_{reg}(U) = \iint_{T \times \Omega} \left( \frac{\iint_{T \times \Omega} w(\mathbf{x}, t, \mathbf{y}, t_1) U_{t_1}(\mathbf{y}) d\mathbf{y} dt_1}{\iint_{T \times \Omega} w(\mathbf{x}, t, \mathbf{y}, t_1) d\mathbf{y} dt_1} - U_t(\mathbf{x}) \right)^2 dxdt \tag{2}$$

where  $w(\mathbf{x}, t, \mathbf{y}, t_1)$  is a similarity measure between two pixels  $\mathbf{x}$  and  $\mathbf{y}$  observed on the frame  $t$  and  $t_1$ . Finally, we define the global cost function that is minimized to estimate the restored image as

$$E(u) = E_{reg}(U) + \lambda \int_0^T E_{data}(U_t) dt \tag{3}$$

where  $\lambda$  is a trade-off parameter between the regularization and the fidelity to data term. The estimation of the noise free ultrasound sequence is obtained through minimizing the cost function given in equation (3). The minimization is done through a gradient descent scheme. Starting from the noisy observation  $U^0 = I$ , the restored image is obtained at the convergence of the sequence  $U_t^k$

$$U_t^{k+1}(\mathbf{x}) = U_t^k(\mathbf{x}) - dt \left[ \frac{\partial E_{reg}}{\partial U_t(\mathbf{x})} + \lambda \left( \frac{1}{U_t(\mathbf{x})} - \frac{1}{\sigma_n^2} \frac{I_t(\mathbf{x})}{U_t(\mathbf{x})^3} \right) \right] \tag{4}$$

In order to decrease the computation cost we will restrict the interaction between pixels to a local neighborhood domain and not the whole image. Therefore, we note :  $\Pi_{\mathbf{x}}$  the spatial neighborhood centered on a pixel  $\mathbf{x}$  and  $T_w$  the size of

the temporal window (the number of frames that interact directly with a given frame). In this case, the derivative of the regularization term is equal to

$$\begin{aligned}
 \frac{\partial E_{reg}}{\partial U_t(\mathbf{x})} &= 2 \left( U_t(\mathbf{x}) - \int_{\Pi_{\mathbf{x}}} \int_{t-T_w}^{t+T_w} \frac{w(\mathbf{x}, t, \mathbf{y}, t_1)}{Z(\mathbf{x}, t)} U_{t_1}(\mathbf{y}) d\mathbf{y} dt_1 \right) \\
 + 2 \int_{\Pi_{\mathbf{x}}} \int_{t-T_w}^{t+T_w} &\left( U_{t_1}(\mathbf{z}) - \int_{\Pi_{\mathbf{z}}} \int_{t_1-T_w}^{t_1+T_w} \frac{w(\mathbf{z}, t_1, \mathbf{y}, t_2)}{Z(\mathbf{z}, t_1)} U_{t_2}(\mathbf{y}) d\mathbf{y} dt_2 \right) \frac{w(\mathbf{z}, t_1, \mathbf{x}, t)}{Z(\mathbf{z}, t_1)} d\mathbf{z} dt_1
 \end{aligned} \tag{5}$$

with  $Z(\mathbf{x}, t) = \int_{\Pi_{\mathbf{x}}} \int_{t-T_w}^{t+T_w} w(\mathbf{x}, t, \mathbf{y}, t_1) d\mathbf{y} dt_1$  the normalization coefficient. Note that such a derivative formulation assumes that the weights are independent from the image  $U$  which is a valid assumption because weights are calculated in the beginning of the process on the noisy observations.

### 3 Weights Computation

In our algorithm we combine spatial and temporal filtering. Therefore, a robust weight definition is crucial to avoid details blurring due to motion. More explicitly, with a good definition of weights there is no need to estimate the motion of the organ. In this work, we will suggest a suitable definition of the weights in the case of multiplicative Rayleigh noise and we will consider a distance based on the spatial information as well as the temporal one. To this end, we use patch based similarity which is more robust to noise than the similarity between intensities at a pixel level. We recall that this measure relies on the assumption that an image is stationary (at least at local scale), and redundant which means that a given patch has several copies in the entire ultrasound sequence. Let us characterize a pixel  $\mathbf{x}$  at the frame  $U_t$  by the set of neighboring pixels in the same frame. Let us consider now the  $L^2$  distance between two patches of size  $(2p + 1) \times (2p + 1)$  centered on  $\mathbf{x}$  and  $\mathbf{y}$  in two different frames  $U_{t_1}$  and  $U_{t_2}$  in the sequence.

$$\begin{aligned}
 d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) &= \sum_{\mathbf{z} \in [-p, p]^2} (I_{t_1}(\mathbf{x} + \mathbf{z}) - I_{t_2}(\mathbf{y} + \mathbf{z}))^2 \\
 &= \sum_{\mathbf{z} \in [-p, p]^2} (U_{t_1}(\mathbf{x} + \mathbf{z})n_{t_1}(\mathbf{x} + \mathbf{z}) - U_{t_2}(\mathbf{y} + \mathbf{z})n_{t_2}(\mathbf{y} + \mathbf{z}))^2
 \end{aligned}$$

In case the observed patches are derived from the same speckle free patch, we can write:

$$d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) = \sum_{\mathbf{z} \in [-p, p]^2} U_{t_1}^2(\mathbf{x} + \mathbf{z})(n_{t_1}(\mathbf{x} + \mathbf{z}) - n_{t_2}(\mathbf{y} + \mathbf{z}))^2$$

This equation shows that contrarily to additive noise, the  $L^2$  distance is not only dependent on noise distribution but also on the image intensity. To overcome this dependence one can compute

$$d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) = \sum_{\mathbf{z} \in [-p, p]^2} \frac{(I_{t_1}(\mathbf{x} + \mathbf{z}) - I_{t_2}(\mathbf{y} + \mathbf{z}))^2}{U_{t_1}^2(\mathbf{x} + \mathbf{z})}$$

To determine the distance  $d_s(\mathbf{x}, t_1, \mathbf{y}, t_2)$ , one needs an estimation of  $U_{t_1}(\mathbf{x})$ . A simple way to do that is to compute the average of  $I_{t_1}$  over the patch around  $\mathbf{x}$ . Furthermore, in order to obtain a symmetric distance with respect to  $t_1$  and  $t_2$  we will consider  $U_{t_1}^2 = U_{t_1} * U_{t_2}$ . Thus, we obtain the following distance definition

$$d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) = \sum_{\mathbf{z} \in [-p, p]^2} \frac{(I_{t_1}(\mathbf{x} + \mathbf{z}) - I_{t_2}(\mathbf{y} + \mathbf{z}))^2}{\tilde{U}_{t_1}(\mathbf{x} + \mathbf{z})\tilde{U}_{t_2}(\mathbf{y} + \mathbf{z})} \tilde{U}_{t_1}(\mathbf{x}) = \sum_{\mathbf{z} \in [-p, p]^2} \frac{I_{t_1}(\mathbf{x} + \mathbf{z})}{(2p + 1)^2}$$

Now if we consider the temporal aspect of the sequence, two similar pixels that belong to the same structure have similar displacement vectors. Thus, in addition to the distance between patches, the variation of pixel intensity over time should be taken into account. To this end, we consider temporal windows of size  $(2t_c + 1)$  and similarly to the spatial distance, we define

$$d_t(\mathbf{x}, \mathbf{y}, t_1, t_2) = \sum_{k=-t_c}^{t_c} \frac{[I_{t_1+k}(\mathbf{x}) - I_{t_2+k}(\mathbf{y})]^2}{\tilde{U}_{t_1+k}(\mathbf{x})\tilde{U}_{t_2+k}(\mathbf{y})} \quad (6)$$

Finally we define the similarity measure between  $\mathbf{x}$  and  $\mathbf{y}$  belonging to two different frames as a decreasing function of the distances  $d_s(\mathbf{x}, \mathbf{y}, t_1, t_2)$  and  $d_t(\mathbf{x}, \mathbf{y}, t_1, t_2)$ . A possible weight definition is

$$w(\mathbf{x}, t_1, \mathbf{y}, t_2) = \exp\left(-\frac{d_s(\mathbf{x}, t_1, \mathbf{y}, t_2)}{2h_s^2}\right) \exp\left(-\frac{d_t(\mathbf{x}, t_1, \mathbf{y}, t_2)}{2h_t^2}\right) \quad (7)$$

$h_s$  and  $h_t$  are parameters that have an impact on the selection degree of pixels that interact together. They have an influence on the smoothness of the final result. To summarize, the proposed weights are designed to take into account the noise properties, mainly the fact that it is multiplicative. With such definition, we restrict interactions between pixels only to similar ones in order to preserve details in each frame. We take also into account motion consistency by computing similarity measure on temporal windows.

## 4 Experimental Results

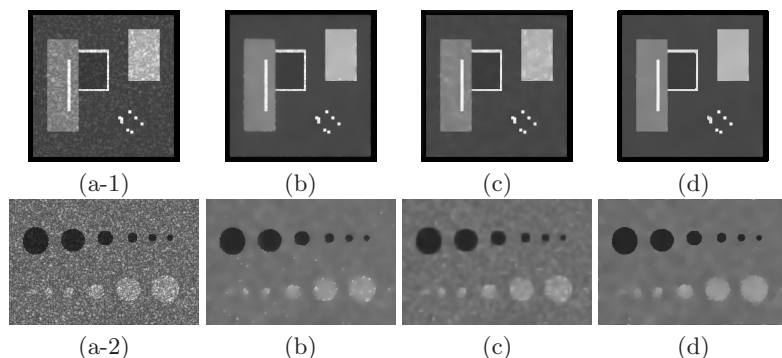
To assess the performance of the proposed method we used both synthetic and real data. In the first experiment, we used synthetic images with artificial speckle. The speckle is simulated by low pass filtering a complex Gaussian random field and computing its magnitude [2]. We compared our algorithm to the anisotropic diffusion algorithm [3] and a wavelet based technique [7] using the Matlab implementation provided by the authors. Regarding anisotropic diffusion the parameter is the number of iteration. For the wavelet based algorithm the parameter are the thresholding coefficient and the neighborhood size were local energy was computed. For the first method we tuned the iteration number to obtain the better possible results (PSNR and visually). The optimal parameters range of the second method was provided by the author in their paper but we slightly

**Table 1.** PSNR values for denoised image corrupted by Speckle

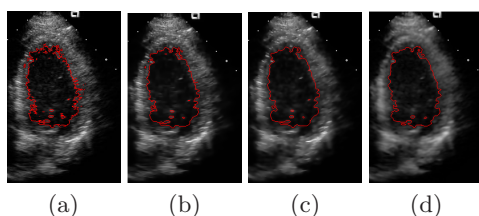
| $\sigma_n$      | Sythetic1 [7] |       |       | Sythetic2 |       |       |
|-----------------|---------------|-------|-------|-----------|-------|-------|
|                 | 1             | 0.5   | 0.25  | 1         | 0.5   | 0.25  |
| Corrupted image | 22.20         | 27.93 | 33.95 | 20.44     | 26.47 | 32.62 |
| OurMethod       | 29.16         | 37.32 | 38.01 | 33.88     | 38.89 | 43.53 |
| SRAD[3]         | 24.88         | 33.09 | 40.95 | 31.55     | 38.36 | 46.15 |
| wav [7]         | 29.14         | 35.48 | 41.69 | 31.06     | 35.85 | 39.75 |

modified them to improve their performance with respect our data. Regarding our algorithm we considered  $11 \times 11$  window size for  $\Pi_x$  while  $\lambda = 0.05$ . For weight computation the parameters setting is:  $h_s = \frac{\sigma_n}{5}$  and a  $3 \times 3$  patch size. We did not consider the temporal aspect for this experiment because the synthetic data are single images and not sequences. For an evaluation we considered the PSNR  $\left[PSNR = 10 \log_{10} \frac{255^2}{MSE}\right]$  of the reconstructed images  $\hat{U}$  that is a function of the mean square error with respect to the ground truth image noted  $\left[MSE = \sum_{t=0}^T \sum_{\mathbf{x} \in \Omega} (U_t(\mathbf{x}) - \hat{U}_t(\mathbf{x}))^2\right]$ .

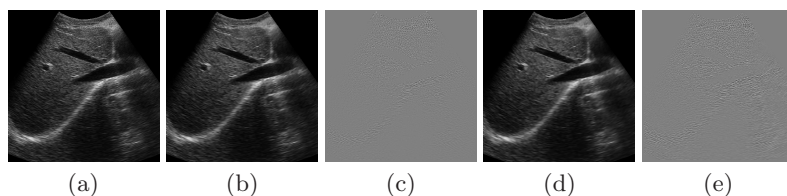
In table (1) we reported the PSNR value for the different methods and with different noise variance. This results show that our method achieve high performance when the noise variance is important. Such a behavior is justified by the fact that we consider large spatial neighborhood ( $\Pi_x$ ) for filtering which reduces considerably the variance inside each region. The anisotropic diffusion algorithm (SRAD)[3] is more local and its numerical scheme is based on interactions between pixels at local scale when compared to our algorithm. Besides, the gradient information is not reliable in case of high level noise. The approach presented in [7] and the choice of the parameter of threshold used in this algorithm to compute the noise free wavelet coefficients is critical to insure the balance between edge preserving and noise suppression. In figure [Fig.(1)] one can see the restoration results for the synthetic image using the different methods and different variance. The method based on the wavelet transform provides images with sharp contours but the noise remains in homogeneous regions. The anisotropic diffusion results in blurry edges while smoothing the homogeneous areas. Our algorithm reaches an optimal balance where the obtained images are completely smoothed inside each region while the edges being sharp. As we stated before we consider large neighborhood size to remove the speckle, while encoding the image structure in our weight definition to preserve image discontinuities and details. This figure shows that our method preserves the image discontinuities and contrast between regions better than the two other algorithms. As far as real data are concerned we considered an ultrasound sequence of the left ventricle on which we applied our speckle removal algorithm. The parameters used for this experiment are: a  $7 \times 7$  window size for  $\Pi_x$ ,  $\lambda = 0.5$  and  $T_w = 5$ . The weight computation was performed using  $3 \times 3$  patch size, a temporal window for comparison of size  $t_c = 2$  and  $h_s = h_t = 0.05$ . In figure [Fig.(2)] we reported the restoration result on one frame of an ultrasound sequence. To evaluate the quality of restoration, we extracted the boundaries of the ventricle, using a level set based technique. The contours obtained using our method are smooth and



**Fig. 1.** (a-1) Image (Synthetic1) corrupted by the speckle  $\sigma_n = 0.5$  (a-2) Image (Synthetic2) corrupted by the speckle  $\sigma_n = 1$  (b) Result using the anisotropic diffusion [3] (c) Result using the wavelet based technique [7] (d) Result using our algorithm



**Fig. 2.** Results of restoration on real ultrasound frames. (a) Observed image (b) Result using algorithm [7] (c) Anisotropic diffusion [3] (d) Results of our algorithm.



**Fig. 3.** Results of filtering real ultrasound frames (a) Observed image (b) Result of our algorithm without temporal component ( $T_w = 0$ ) (frame by frame filtering) (c) Residual obtained with our algorithm without temporal component ( $T_w = 0$ ) (frame by frame filtering)(d) Result of our algorithm using temporal filtering ( $T_w = 2$ ) (e) Residual obtained with our algorithm using temporal filtering ( $T_w = 2$ )

the structure being present in the image were preserved. The contour extracted in the frame restored using our approach is similar to the one associated to the anisotropic diffusion, but this method yields piecewise smooth images. Regarding the wavelet based technique, it produces image with sharp details but the contours are not smooth. To evaluate the impact of the time component we compared the algorithm performance by processing the whole sequence using

$T_w = 2$  and by processing it frame by frame ( $T_w = 0$ ). and the obtained results are shown in figure [Fig.(3)]. One can see that using a temporal window is more efficient for speckle reduction and this is illustrated by the residual images [Fig.(3-c),Fig.(3-e)](difference between the filtered image and the observed one). Besides, we notice that considering the time component doesn't introduce a blur on the images because our weight definition is robust enough to compensate the effect of motion. To conclude we can say that the experimental results show that our method can deal with correlated speckle even though we make the assumption of independence between pixels. We have to point out that the proposed speckle removal approach is flexible with the selection of the scale of interaction between pixels contrarily to the PDE based technique presented in [3].

## 5 Discussion

In this paper we proposed a new regularization technique for ultrasound sequences. The regularization term was based on weight definition adapted to the Speckle to impose some consistency with the observed structure in the image. We introduced also a data term that is coherent with the speckle distribution A future direction to improve our work will be considering anisotropic neighborhood shapes as well as variable size ones. This will ensure a better interaction between pixels lying on the same structure. Considering the fact that speckle is a correlated signal is also a direction that could bring further improvement to the method.

## References

1. Wagner, R., Smith, S., Sandrik, J.M., Lopez, H.: Statistics of speckle in ultrasound b-scans. *IEEE Trans. on Sonics and Ultrasonics* 30(3), 156–163 (1983)
2. Achim, A., Bezerianos, A., Tsakalides, P.: Novel bayesian multiscale method for speckle removal in medical ultrasound images. *IEEE trans. on Medical Imaging* 20, 772–783 (2001)
3. Yu, Y., Acton, S.: Speckle reducing anisotropic diffusion. *IEEE trans. on Image Processing* 11(11), 1260–1270 (2002)
4. Abbott, J.G., Thurstone, F.L.: Acoustic speckle: Theory and experimental analysis. *Ultrason. Imag.* 1, 303–324 (1979)
5. Markosien, L.: A motion-compensated filter for ultrasound image sequences. Technical report, Providence, RI, USA (1996)
6. Angelini, E., Laine, A., Takuma, S., Holmos, J., Homma, S.: Lv volume quantification via spatiotemporal analysis of real-time 3-D echocardiography. *IEEE trans. on Medical Imaging* 20(6), 457–469 (2001)
7. Pizurica, A., Philips, W., Lemahieu, I., Acheroy, M.: A versatile wavelet domain noise filtration technique for medical imaging. *IEEE trans. on Medical Imaging* 22(3), 323–331 (2003)
8. Azzabou, N., Paragios, N., Guichard, F., Cao, F.: Variable bandwidth image denoising using image-based noise models. In: *CVPR*, pp. 1–7 (2007)