Motion Estimation in Ultrasound B-mode Images Using a Maximum Likelihood Matching Criterion.

Isaac Kokkinidis† , Michael G. Strintzis† , N. Maglaveras ‡

† Department of Electrical and Computer Engineering,

‡ Department of Medical informatics,

University of Thessaloniki,

Thessaloniki 54006, GREECE

Email: isaak@eng.auth.gr

Abstract

An optimal method for estimating motion in ultrasound B-mode images is presented. Such estimation is often needed for diagnostic purposes to distinguish regions of the image with differing motion. The block matching technique is adapted to the nature and statistics of these images, using maximum likelihood (ML) estimation. In application to echocardiographic images, this approach is seen to give better results than standard block matching, without significantly increasing the computational complexity of the algorithm.

Keywords: *ultrasound, motion estimation.*

1 Introduction

Motion estimation is useful in medical imaging as a diagnostic tool, as pathological cases are often identified using motion determination. For example, the use of motion estimation to tumor detection in echocardiograms has been described in [1].

The most widely used method for motion estimation is block-matching. This technique has been proven to be very efficient for the compression of image sequences, and its use has been adopted by digital television transmission standards, such as MPEG. However, in ultrasound imaging, due to the poor quality and low contrast of the displayed output, and also the non-Gaussian statistics encountered, the classical block matching algorithm does not perform well, producing block artifacts and unacceptable for diagnostic use motion estimation results.

In this work, novel matching criteria, optimal for ultrasound images are applied. The noise speckles in echo images are modeled as a multiplicative Rayleigh process in B-mode data, and alternately as a signal-dependent Gaussian process in displayed images [2].

2 Optimal block matching

According to the block matching technique, each frame is segmented into small orthogonal blocks, on which a search is performed to find the best matching block of the previous in time frame (backward prediction). Let $\mathbf{b}_i$ be the blocks obtained by segmenting the current frame I_t. These blocks are considered to be on a grid $G(I_t)$, non-overlapping, and their union covers the whole image I_t. We denote by $\mathbf{a}_i$ the matching to $\mathbf{b}_i$ blocks of the previous frame I_{t-1}. These blocks are not snapped to a grid, and they no longer need to be non-overlapping, neither their union covers the image I_{t-1}.

The block matching (BM) criterion is usually the minimization of a norm $\mathcal{L}$ of the difference of the luminance values of the pixels of these blocks. Specifically the problem is:

$$\begin{gathered}\forall i \in G(I_t), \quad \forall (dx, dy) \in W, \text{find the vector} \\ v_i = [dx, dy]^T \quad \text{that minimizes the norm:} \\ \mathcal{L}_{\forall_{\text{pels}} \in w}[\mathbf{b}_i(x,y), \mathbf{a}_i(x+dx, y+dy)]\end{gathered}$$

where w is the block area, W is the search area, which is a square region around the reference block's coordinates x, y, and dx, dy are the horizontal and vertical displacements of the reference block.

Our goal is to estimate the vector field that best represents the motion parameters from the previous to the current frame. Considering the observations I_{t-1} and I_t of these two images, the maximum likelihood estimator of the vector field $\mathbf{v}$, is determined by the maximization of the conditional probability:

$$\max_{\mathbf{v}}\{P(I_{t-1}, I_t|\mathbf{v})\} \quad \text{for} \quad \mathbf{v} = \hat{\mathbf{v}}_{\text{ML}} \tag{1}$$

Which, using the Bayes' formula, is seen to be equivalent to [3]:

$$\max_{\mathbf{v}}\{P(I_{t-1}|I_t, \mathbf{v})\} \quad \text{for} \quad \mathbf{v} = \hat{\mathbf{v}}_{\text{ML}} \tag{2}$$

With the above block formulation, (2) can be broken into a product of block-based conditional densities:

$$\prod_{i \in G(I_t)} \left\{ \max_{v_i}[f_{\mathbf{a}_i}(\mathbf{a}_i|\mathbf{b}_i, v_i)] \right\} \tag{3}$$

with which we can search for block maxima in the previous image.

3 Matching criterion derivation

3.1 Multiplicative noise model

Let the intensities of the pixels of frame I_{t-1}, I_t be denoted by a, b, respectively. According to the multiplicative noise model for *matching* pixels the ratio of the observed signals a/b may be modeled as a Rayleigh process n with expected value equal to 1, which is statistically independent of b. In other words, we assume that b contains the same information as a, but it is corrupted by multiplicative Rayleigh noise. The conditional density $f_a(a|b,v)$ for *matching* pixels can then be expressed in terms of the random variable $n = a/b$:

$$a = nb \Rightarrow f_a(a|b,v) = \frac{1}{b} f_n\left(\frac{a}{b}\right) \tag{4}$$

The probability density function of the noise random variable n is written:

$$f_n(n) = \frac{\pi n}{2} \exp\left[-\frac{\pi n^2}{4}\right], \quad n > 0 \tag{5}$$

Extending the above to blocks, we consider the vector random variables $\mathbf{a}_i, \mathbf{b}_i, \mathbf{n}_i$ for $i \in G(I_t)$, which are defined as:

$$\mathbf{a}_i = [a_{i1}, a_{i2}, \ldots, a_{ik}]^T \tag{6}$$

$$\mathbf{b}_i = [b_{i1}, b_{i2}, \ldots, b_{ik}]^T \tag{7}$$

$$\mathbf{n}_i = [n_{i1}, n_{i2}, \ldots, n_{ik}]^T = \left[\frac{a_{i1}}{b_{i1}}, \frac{a_{i2}}{b_{i2}}, \ldots, \frac{a_{ik}}{b_{ik}}\right]^T \tag{8}$$

where k is the number of pixels per block.

Assuming that the pixel values in a block w are independent random variables, we have for $i \in G(I_t)$ [3]:

$$f_{\mathbf{n}_i}(\mathbf{n}_i) = \prod_{j=1}^{k} \left\{ \frac{\pi n_{ij}}{2} \exp\left[-\frac{\pi n_{ij}^2}{4}\right] \right\} \tag{9}$$

where f_{n_i} are given by (5).

The maximum likelihood criterion for each block from (3), (4) and (9) is written:

$$\max_{v_i} \left\{ \prod_{j=1}^{k} \left[\frac{1}{b_{ij}} f_{n_{ij}} \left(\frac{a_{ij}}{b_{ij}} \right) \right] \right\}, \quad \text{hence :}$$

$$\max_{v_i} \left\{ \prod_{j=1}^{k} \left[\frac{a_{ij}}{b_{ij}^2} \exp\left(-\frac{\pi a_{ij}^2}{4 b_{ij}^2} \right) \right] \right\} \tag{10}$$

where the maximization is achieved by searching within a window on the previous image. In (10), a_{ij} are the intensities of the frame I_{t-1}, and b_{ij} the corresponding pixels of I_t if motion v_i is assumed.

3.2 Signal-dependent noise model

An alternate model for ultrasound images is $a = b + b^{1/2} n$ [2], where the noise n is a zero-mean Gaussian random variable:

$$f_n(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{n^2}{2\sigma^2}\right] \tag{11}$$

Using the same conventions as in subsection 3.1, i.e. for a, b *matching* pixels of I_{t-1}, I_t, the conditional probability in (3) can be written as:

$$f_a(a|b,v) = \frac{1}{\sqrt{b}} f_n\left(\frac{a-b}{\sqrt{b}}\right) \tag{12}$$

Extending to blocks and using (6), (7) and (8), we conclude with the matching criterion:

$$\max_{v_i} \left\{ \prod_{j=1}^{k} \left[\frac{1}{\sqrt{b_{ij}}} \exp\left(-\frac{(a_{ij} - b_{ij})^2}{2\sigma^2 b_{ij}} \right) \right] \right\} \tag{13}$$

frame	ML §3.1	ML §3.2	Classical BM
1	20.74	20.23	20.16
2	20.77	19.47	18.48
3	20.57	19.80	19.02
4	20.76	20.34	19.24
5	22.18	20.98	20.37

Table 1: PSNR in dB of the estimated images.

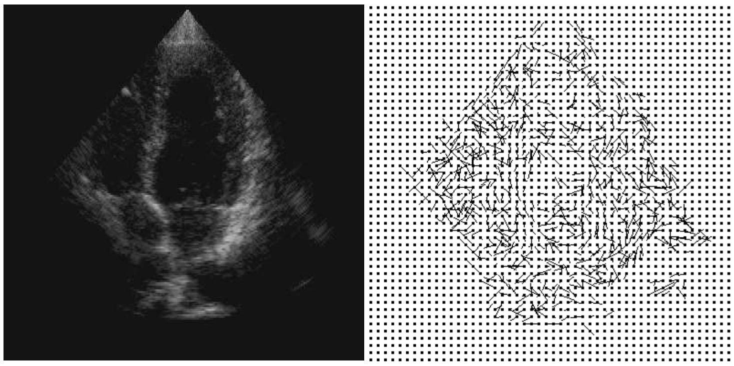

Figure 1: A sample frame and the vector field obtained by the algorithm described in §3.1.

4 Results and Conclusions

The above formulas were used on sets of images acquired with the trans-esophageal technique using digital equipment. The block size used for the motion estimation was 8×8 pixels, and the searching area was taken as a square of 15×15 pixels centered at the reference block's coordinates (x, y). Half pixel motion vectors were found.

In Table 1 are shown typical results of motion estimation in these sequences, using the methods of sections 3.1 and 3.2 above, compared with standard block matching. The first method presented (that in section 3.1), is seen to clearly perform the best.

Acknowledgement

This work was supported by the PENED 91ED838 project administered by the General Secretariat of Research and Technology of Greece.

References

[1] C. Kontogeorgakis, M. G. Strintzis, N. Maglaveras, I. Kokkinidis, "Tumor Detection in Ultrasound B-mode Images Through Motion Estimation Using a Texture Detection Algorithm", *Computers in Cardiology 1994*, September 1994.

[2] C. Kotropoulos, X. Magnisalis, I. Pitas, M. G. Strintzis, "Nonlinear Ultrasonic Image Processing Based on Signal-Adaptive Filters and Self-Organizing Neural Networks", *IEEE Trans. on Image Processing*, Vol. 3, No. 1, pp 65-78, January 1994.

[3] A. Papoulis, "*Probability, Random Variables, and Stochastic Processes*", McGraw-Hill Inc, 1991.