Ultrasound in Med. & Biol., Vol. ■, No. ■, pp. 1–19, 2015

0301-5629/$ - see front matter

ELSEVIER

http://dx.doi.org/10.1016/j.ultrasmedbio.2015.07.022

● *Review Article*

ESTIMATION OF LARGE-SCALE ORGAN MOTION IN B-MODE ULTRASOUND IMAGE SEQUENCES: A SURVEY

Valeria De Luca, Gábor Székely, and Christine Tanner
Computer Vision Laboratory, ETH Zurich, Zurich, Switzerland

(*Received* 11 *August* 2014; *revised* 13 *June* 2015; *in final form* 16 *July* 2015)

Abstract—Reviewed here are methods developed for following (*i.e.*, tracking) structures in medical B-mode ultrasound time sequences during large-scale motion. The resulting motion estimation problem and its key components are defined. The main tracking approaches are described, and their strengths and weaknesses are discussed. Existing motion estimation methods, tested on multiple *in vivo* sequences, are categorized with respect to their clinical applications, namely, cardiac, respiratory and muscular motion. A large number of works in this field had to be discarded as thorough validation of the results was missing. The remaining relevant works identified indicate the possibility of reaching an average tracking accuracy up to 1–2 mm. Real-time performance can be achieved using several methods. Yet only very few of these have progressed to clinical practice. The latest trends include incorporation of complementary and prior information. Advances are expected from common evaluation databases and 4-D ultrasound scanning technologies. (E-mail: vdeluca@vision.ee.ethz.ch) © 2015 World Federation for Ultrasound in Medicine & Biology.

Key Words: **Tracking, Ultrasound imaging, Motion estimation, Displacement estimation, Image registration, Organ motion, Cardiac motion, Respiratory motion, Ultrasound-guided therapy, Ultrasound-guided intervention, Review.**

INTRODUCTION

Ultrasound (US) imaging is a well-established and widely used medical imaging technique. Echographic systems are inexpensive, easily accessible and relatively small and mobile and do not require dedicated rooms. US is a safe image modality as it does not use ionizing radiation. In addition, US imaging is real time and can generate image sequences at a high frame rate. Most available systems are using 2-D imaging. As an emerging technology, 3-D US systems allow real-time volumetric imaging, which is particularly appealing for direct visualization of 3-D anatomy and motion analysis. Yet high temporal resolution comes currently at the expense of reduced spatial resolution. In clinical practice, US is commonly used for detection, diagnosis and follow-up, especially for cardiac screening (echocardiography), abdominal imaging, obstetrics, vascular imaging and further investigation of breast cancer (Fenster et al. 2001).

Address correspondence to: Valeria De Luca, Computer Vision Laboratory, ETH Zurich, 8092 Zurich, Switzerland. E-mail: vdeluca@vision.ee.ethz.ch

The capability of visualizing soft tissue deformations and organ displacements in real time makes US an appealing choice for medical applications that require tissue tracking and analysis of the estimated motion. These include strain rate imaging (Yip et al. 2003), ventricular deformation analysis (Frangi et al. 2001) and motion compensation in image-guided interventions (Cleary and Peters 2010) and therapy (Fontanarosa et al. 2015; Schwaab et al. 2014). Recent technical developments in radiation therapy, such as intensity-modulated radiation therapy and high-intensity focused ultrasound, can deposit highly conformal dose distributions into the tissues. Therefore, compensation for any movement in the treatment region, like respiratory motion in the abdomen, is vital. This requires a tracking accuracy in the range of millimeters and real-time capability, that is, at least as fast as the image acquisition rate (Keall et al. 2006; Shirato et al. 2007).

Compared with tomographic imaging, such as magnetic resonance imaging (MRI) and computed tomography (CT), US images are harder to understand and interpret. They are often limited by small acoustic windows, and their quality is affected by several types of noise and artifacts. These include shadows and mirroring, which are due mainly to highly attenuating structures like bones and/or strongly reflecting interfaces. These artifacts are useful for some applications to understand the tissue composition, for example, shadowing below a breast lesion. Yet they are not stable in the presence of moving tissue, as they depend on the tissue composition along the US beam and, hence, represent an obstacle in estimating motion.

To track (*i.e.*, follow) structures on the (US) image sequence over time, the essential task is to estimate the motion (*i.e.*, displacement) of the structures. To achieve this, the spatial correspondence between structures on different frames has to be determined. The resulting displacement vectors between consecutive frames can be concatenated to provide the motion trajectory of the tracked structure. In the clinic, tracking information provides feedback about the actual position to allow for motion compensation in therapy and surgery guidance, and for quantitative physiopathologic analysis. In this article, we consider only large-scale motions, such as the contraction of the myocardium (~10 mm [Andreu et al. 2012]), the motion induced in the liver by respiration (10–20 mm [Keall et al. 2006]) and muscle fascicle movement (~55° and 20 mm [Gillett et al. 2013]).

There are typically two kinds of US data, radiofrequency (RF) and in-phase and quadrature (IQ) data. Different modes of US are available, including brightness (B-mode) and harmonic mode, contrast imaging and Doppler ultrasonography. Their use depends on the specific clinical application. For large motion estimation, B-mode imaging is the most commonly employed modality, and hence we restrict our review to B-mode imaging.

In this article, we review tracking applications and hence we consider only the contributions that were evaluated on 2-D and/or 3-D sequences of more than two consecutive images that were continuously acquired at the same spatial location. Furthermore, to keep the review concise and most relevant, we focused on the methodologies, which were validated on sequences from real tissues undergoing real motion (excluding the ones validated only on synthetic, simulated and phantom images) and whose results were quantitatively evaluated with respect to manual annotation of (clinical) experts. We selected only the studies in which the acquisition and use of data from human or animal subjects were approved by an ethics committee or institutional review board, where applicable, and in which informed consent by each human study participant was received. The inclusion was based on written confirmation when this information was not explicitly stated.

A literature search was performed looking for works published in the past 20 years, until January 2015. We searched for the words *ultrasound*, *motion*, *tracking* and *registration* using Google Scholar, NCBI PubMed, IEEE Xplore, ScienceDirect and SpringerLink search engines, and looked through the references in the found articles. The criteria for selecting the contributions included in this review are listed in Figure 1.

Image classification, compounding, segmentation, blood flow tracking, registration of non-consecutive images and multimodal approaches are not reviewed here. See the works by Rao et al. (2006), Solberg et al. (2007), Hansen et al. (2010) and Zahiri-Azar et al. (2011) for beam steering and angular compounding; Noble and Boukerroui (2006) for ultrasound segmentation; and Bjaerum et al. (2002) and Jensen (1996) for blood flow tracking. We do not review strain imaging, elastography and time delay estimation methods, as they focus on determining the mechanical properties of soft tissue, usually cover small-scale displacements and are based on RF or IQ data. Furthermore, several comprehensive works and reviews on the topic are already available in the literature; see Ophir et al. (1991, 1999), Nightingale et al. (2002), Greenleaf et al. (2003), Zahiri-Azar and Salcudean (2006), Parker et al. (2011) and Gennisson et al. (2013).

The article is structured as follows. Under Methodology Approaches, we define the tracking problem in US sequences and give an overview of the main US tracking components; we also list the main motion estimation strategies and classify them according to the image features used. Under Validation Approaches, we discuss the importance of US tracking validation and the most common performance measures. Under Clinical Applications, we describe the most significant contributions in the context of clinical applications, and we present the state-of-the-art for the most relevant clinical areas. These are, for this review, cardiac, abdominal and muscular motions. The last section concludes the survey with a global assessment of the current state of research and presents possible future research directions.

METHODOLOGY APPROACHES

In a clinical scenario, US sequences are often acquired for estimating the motion of anatomic structures. The movement of a structure is described by motion vectors, obtained by finding the spatial correspondence between structures in consecutive images. Image registration, which is the process of determining spatial correspondence between two or more images (*i.e.*, spatial transformation for aligning the moving image[s] with the fixed image) rather than individual structures, is often locally employed for tracking structures. Hence we describe the main components of motion estimation algorithms in the framework of image registration and highlight differences. Note that, in comparison to image registration, tracking algorithms also need to be fast