A DYNAMIC PROGRAMMING APPROACH FOR AUTOMATIC STRIDE DETECTION AND SEGMENTATION IN ACOUSTIC EMISSION FROM THE KNEE

Costas Yiallourides[1], *Victoria Manning-Eid*[2], *Alastair H. Moore*[1] *and Patrick A. Naylor*[1]

[1]Department of Electrical and Electronic Engineering, Imperial College London, UK
[2]MSk Lab, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, UK
{costas.yiallourides08, v.manning, alastair.h.moore, p.naylor}@imperial.ac.uk

ABSTRACT

We study the acquisition and analysis of sounds generated by the knee during walking with particular focus on the effects due to osteoarthritis. Reliable contact instant estimation is essential for stride synchronous analysis. We present a dynamic programming based algorithm for automatic estimation of both the initial contact instants (ICIs) and last contact instants (LCIs) of the foot to the floor. The technique is designed for acoustic signals sensed at the patella of the knee. It uses the phase-slope function to generate a set of candidates and then finds the most likely ones by minimizing a cost function that we define. ICIs are identified with an RMS error of 13.0% for healthy and 14.6% for osteoarthritic knees and LCIs with an RMS error of 16.0% and 17.0% respectively.

Index Terms— acoustic emission, dynamic programming, knee osteoarthritis, contact instants

1. INTRODUCTION

Osteoarthritis (OA) is the most common and disabling of all musculoskeletal diseases. It occurs most frequently in the knee, affecting as many as 1 in 5 adults over the age of 45 [1]. Symptoms include pain, stiffness and swelling, greatly affecting quality of life. OA is associated with increased mechanical wear, such as through older age and high body weight [2]. No cure exists. Treatments aim to manage symptoms through lifestyle modification, physio- and pharmacological- therapy [1]. More severe cases require a total knee replacement.

Diagnosis of OA relies on a combination of patient reported symptoms and medical imaging (X-ray, Magnetic Resonance Imaging (MRI), ultrasound) of cartilage and subchondral bone degradation. However, sensitivity and accessibility (due to high cost and associated risks) of current imaging methods in early disease is poor. They capture the knee statically or during passive motion and not during functional Open Chain Activity (OCA) i.e. when the foot leaves and makes contact with the ground. Thus at the time of diagnosis, OA is already an end-stage disease, and understanding of its aetiology and progression is still limited. New imaging methods which are sensitive, inexpensive and risk-free are required for the early detection of pre-clinical OA to facilitate effective intervention toward disease management and prevention.

Using body sounds for diagnostic purposes is well documented. Sounds produced by the knee likely depend on the angle of the bones, severity of degradation, lubrication and wear of cartilage. Blodgett (1902) observed increased knee sound with subject age [3] and Bircher (1913) reported distinctive sounds associated with type of knee meniscal injury [4, 5]. Chu et al. [6–9] reported that, unlike clinically healthy knees, the spectral activity (recorded during active motion) of pathological knees spanned the entire audible frequency range and signal acoustic power increased with severity of cartilage damage. With the development of Vibroarthrography (VAG), signal processing algorithms have been proposed to extract relevant information from the non-stationary VAG signals and classify them according to knee pathology [10–13]. VAG relies on accelerometer sensors that are sensitive below frequencies of 1 kHz to pick up mechanical vibrations generated by the movement of articular surfaces. Recently, Mascaro et al. explored the use of Acoustic Emission (AE) as a biomarker for assessing knee joints [14]. By developing a system consisting of electro-goniometers to track joint angle and 2 piezoelectric contact sensors operating in the ultrasonic domain to record AE signals, they used the number of AE bursts to differentiate between healthy and OA knees. In a follow up study by Shark et al. [15,16], by extracting the peak magnitude value and the average signal level of each AE burst and performing Principal Component Analysis, they demonstrated that AE measures of healthy and OA knees form separate clusters. They concluded that OA knees produce substantially more AE events with higher peak magnitude and average signal level values than healthy knees [17].

However, no work has yet explored knee AE acquired during functional OCA such as walking, essential for understanding OA development and progression. Therefore, this paper aims to detect and segment the walking strides in an AE signal. This is fundamental for the acoustic analysis that will enable intermittent sounds to be discriminated from those that occur in every stride (stride-synchronous analysis). Stride synchronization would also enable inter-stride signal aver-

aging to improve detection of quasi-periodic acoustic events with poor SNR as well as the extraction of more targeted features for classification. This paper presents a dynamic programming (DP) algorithm for estimating both the ICIs and the LCIs, the time instants enabling segmentation. Section 2 describes the data acquisition system and the testing protocol. Section 3 presents the DP algorithm. Experiments and results are given in section 4. Conclusions are drawn in Section 5.

2. DATA ACQUISITION

Adults with clinical knee OA (1986 ACR) and reporting no previous knee pain (>2 weeks) were recruited from a London orthopaedic clinic. Knees were classified as: 1) OA, 2) healthy. Exclusion criteria were: aged <18 years, previous surgery, unable to provide consent. AE were acquired (rate 44-48 kHz) using a contact microphone (Basik Pro, Schertler, 20 Hz to 20 kHz) attached over the patella, supported by a digital preamplifier (RME Babyface; PreSonus DigiMax LT), during walking on a specialised treadmill instrumented with force plates. Force plates recorded kinetic and spatiotemporal gait characteristics during stance phase (foot in contact with ground), including ground reaction force at heel strike (ICI), mid-stance and push-off (LCI) and stride length and width.

The assessment commenced with a 5 minute warm-up then walking at 1) progressive speeds on a flat level (starting at 4 km/h) 2) fixed speed up a progressive incline (4 km/h, maximum incline 20%), 3) progressive speeds on a fixed decline. The speeds and inclines were subject dependent. In the evaluation database we have a number of recordings for each knee. These correspond to different assessment stages.

3. ALGORITHM DESCRIPTION

The proposed algorithm operates in two stages. The first part generates a set of candidate instances using the negative-going zero crossings (NZC) of the energy weighted Group Delay (GD) [18]. The second part employs DP to refine the selection by minimising a cost function that we define.

3.1. Candidate Generation

Suppose that sample n of an AE signal $s(n)$ can be predicted from linear combinations of past samples weighted by some predictor coefficients. We can use autocorrelation Linear Predictive Coding (LPC) to find these predictors and apply the inverse filter formed by them to $s(n)$ [19]. This yields the Linear prediction (LP) residual $u(r)$, an impulsive excitation signal shown in Fig. 1(c). These impulsive events indicate particular points in the gait cycle that are to some extent 'unpredictable' from the previous samples. Their location can be accurately estimated by GD. In our algorithm the energy weighted formulation of the GD is used due to its higher performance and lower computational cost compared to other

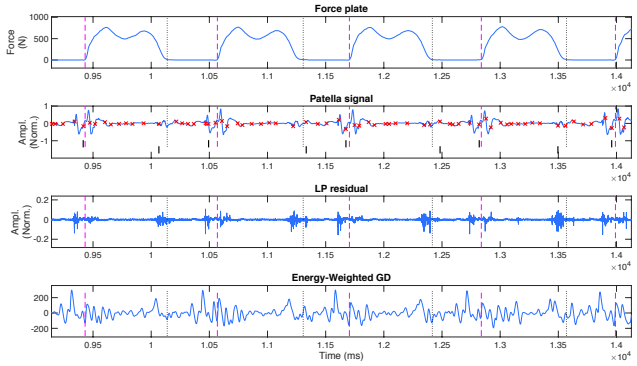

Fig. 1. (a) Force plate signal with identified ICIs (dashed lines) and LCIs (dotted lines) (b) OA knee patella recording with candidates (crosses) and reference CIs (c) LP residual with reference CIs (d) Energy-Weighted GD with reference CIs. Ticks in (b) correspond to selections made by the DP

GD methods, [18], given by:

$$d_{EW}(r) = \frac{\sum_{n=0}^{N-1} n x_r^2(n)}{\sum_{n=0}^{N-1} x_r^2(n)} \quad (1)$$

where $x_r(n) = w(n)u(n+r)$ is an N-sample windowed segment of $u(r)$ beginning at sample r, for $n = 0, 1, ...N-1$. The choice of window size is a compromise as shown in [18]. If it spans the entire stride length then there would be a single NZC per stride corresponding to the strongest excitation event. In this case, inaccurate detection occurs since the highest peak in any stride, as represented in $u(r)$, does not always correspond to only the ICI or the LCI; it can be both, even in the same patient. If it's larger than one stride, more than one event may be included resulting in an NZC at the wrong position. Identification of both ICIs and LCIs requires a smaller window but when it is much smaller than the stride length it is likely that some windows will contain no impulsive events, giving rise to spurious NZC. An increased number of candidates is not problematic however, since these spurious candidates will not be selected by the algorithm if they are, as is the intention, assigned a high cost. An analysis of a number of window sizes and LPC orders is given in Section 4.

3.2. Dynamic Programming

Given the set of candidate contact instants (CIs), the next task is to retain only those which, taken together, best fit the constraints of the system and our prior understanding of the structure of $u(r)$. Each constraint is incorporated by assigning an associated cost to each candidate. The set of CIs which jointly minimise the total cost (2) is then solved using DP [20].

$$\min_{L} \sum_{k=1}^{|L|} \theta_{slope} C_{slope}(k) + \theta_{pd} C_{pd}(k) + \theta_{er} C_{er}(k) \quad (2)$$

where L is the subset selected by the DP, $|L|$ is the number of selections and k is the index of the selection in L. The composing terms (described in the equations to follow) have a range of values from 0 to 1. We use q_{k-1}, q_k and q_{k+1} defined as the sample instances of the candidates $k-1$, k and $k+1$ respectively which depend on L. Our DP algorithm maintains two possible routes for q_k at the same time and considers all candidates to be ICIs for the first route and LCIs for the second. The difference lies in the distribution from which the Energy Ratio cost is derived, (9a) and (9b). At each step two candidates are kept in memory for each route: (a) q_{k-1}, the previous candidate in path (primary path) and (b) the candidate between q_{k-1} and q_k (secondary path), selected based on the swing-to-stance cost (10). At the end of the process the path with the smallest accumulated cost from both routes is selected as the primary. Its elements are labeled as ICIs or LCIs based only on their total Energy Ratio cost depending on which is smallest. Given the label, the appropriate secondary path is chosen.

1) ***Slope Deviation Cost:*** The slope in the vicinity of the NZC in d_{EW} for a clear impulse is -1 by definition. The events in $u(r)$ are not true impulses and hence the slope will deviate from unity. However, we have found that candidates corresponding to true CIs have a slope much closer to -1 than others. This gives us a way of discriminating true events from spurious NZC where the slope is nearly flat. We define

$$C_{slope}(k) = 1 + m(q_k) \tag{3}$$

$$m(q_k) = \frac{1}{l}\left[d_{EW}(q_k + \frac{l}{2}) - d_{EW}(q_k - \frac{l}{2})\right] \tag{4}$$

where l is the length of window in samples centered on the NZC. From our tests we have found 3 ms to be a satisfactory choice for the window duration.

2) ***Period Deviation Cost:*** It is based on the assumption of relatively constant stride period and is defined as:

$$C_{pd}(k) = \begin{cases} \exp(-\frac{\alpha}{(\epsilon - \Delta_p)}) & \text{if } \Delta_p < \epsilon \\ 0 & \text{if } \Delta_p \geq \epsilon \end{cases} \tag{5}$$

$$\Delta_p = \frac{\min[(q_k - q_{k-1}), (q_{k+1} - q_k)]}{\max[(q_k - q_{k-1}), (q_{k+1} - q_k)]}. \tag{6}$$

The cost increases nonlinearly with period ratio (Δ_p) and parameter α controls the rate of increase of the cost: the smaller the value the faster the increment. We have set α to 0.001. The tolerance factor ϵ is defined as the variability in stride duration during normal walking. Its value is determined from the cumulative distribution function of the period ratio as calculated based on the reference instants from the force plate data of the database. It should be dependent on the choice of the GD window size. In the evaluation we have set it to 0.995.

3) ***Energy Ratio Cost:*** Let the energy ratio between q_k and q_{k-1} CI candidate be

$$R_{q_k, q_{k-1}} = \frac{\min[E(q_k), E(q_{k-1})]}{\max[E(q_k), E(q_{k-1})]} \tag{7}$$

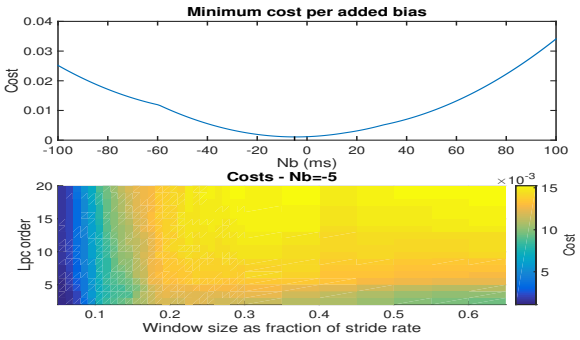

Fig. 2. (a) Minimum cost per N_b (b) Detailed cost variations of N_b = -5 ms for all orders and window sizes

with

$$E(q_k) = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} s^2(n). \tag{8}$$

We define the energy ratio cost as:

$$C_{er,I}(k) = 1 - p_I(f(R_{q_k, q_{k-1}})) \tag{9a}$$

$$C_{er,L}(k) = 1 - p_L(f(R_{q_k, q_{k-1}})) \tag{9b}$$

where p_I and p_L are the probability density of the energy ratio of ICIs and LCIs respectively, trained using ground truth CIs from the force plate signals, and N is the size in samples of a window centered on the candidate. This cost effectively means that we penalise candidates that do not have consistent energies as they are likely not to be true CIs. In the evaluation we found that $N = 0.2fs$ gives good energy consistency.

4) ***Swing-to-Stance Cost:*** It is a function of q_k and q_{k-1} and considers all candidates between the two as possible selections for the secondary path. It is not part of the overall cost function but is rather used to decide on the secondary path based on the selections of the primary. It is defined as:

$$C_{swst}(k) = 1 - p_{swst}(swing/stance) \tag{10}$$

where p_{swst} is the probability density of the swing to stance distribution trained using ground truth CIs. For ICI route, swing is the sample difference between q_k and a viable candidate and stance is the sample difference between the viable candidate and q_{k-1} (reversely defined for the LCI route).

4. EXPERIMENTS AND RESULTS

A database of 74 patella recordings from 11 healthy (56 recordings) and 4 OA knees (18 recordings) was used to evaluate the algorithm against ground truth CIs obtained using a threshold of $10N$ on the corresponding synchronised force plate signals. An estimated ICI is assigned to the j-th stride if it lies in the interval defined, for a reference ICI (γ_j), as $(\gamma_{j-1} + \gamma_j)/2 \leq n < (\gamma_j + \gamma_{j+1})/2$ for each element in

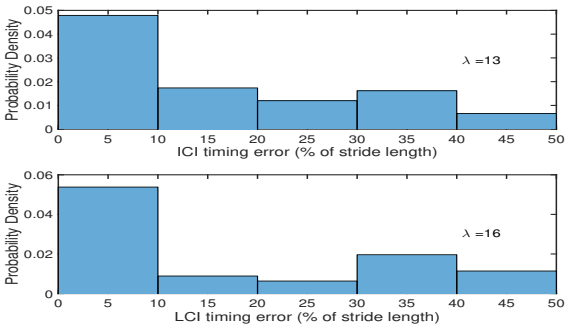

Fig. 3. ICI and LCI timing errors for healthy knees

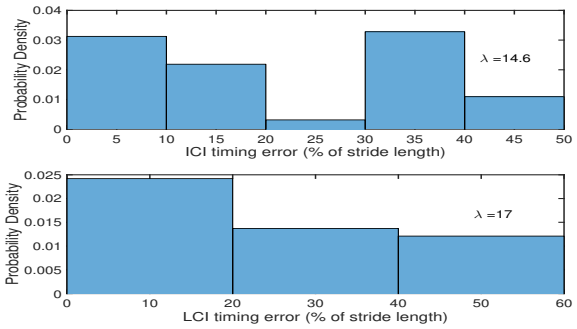

Fig. 4. ICI and LCI timing errors for OA knees

the set $\mathbf{\Gamma} = [\gamma_1, \gamma_2, ..., \gamma_M]$ with $j = 2, 3...M$-1. Similarly, an estimated LCI is assigned to the j-th stride if it lies in the interval defined from the corresponding set of reference LCIs.

The algorithm's performance is assed based on four metrics: 1) identification rate (IDR) - percentage of strides where only a single ICI is detected, 2) miss rate (MR) - percentage of strides where no ICI is detected, 3) false alarm rate (FAR) - percentage of strides where more than one ICI is detected and 4) identification error λ - standard deviation of the timing error (in % of the stride length) between the true and the estimated ICI. Similar reasoning applies for LCIs. Tests were conducted on OA and healthy knees together and using only a single recording from each knee at a speed of 4 km/h on a flat level. Results are reported separately for comparison. The leave-one-out technique was used to generate the empirical distributions in (9a), (9b) and (10) based on the ground truth CIs from all recordings except one which was used for evaluation. Other recordings from the knee under assessment were excluded from the distribution calculation. In all tests we used autocorrelation LPC with 50 ms analysis frames overlapped by 50% to obtain $u(r)$ and the weights of (2) have been empirically determined as $[\theta_{slope}, \theta_{pd}, \theta_{er}] = [0.5, 0.5, 0.5]$.

Evaluation was performed on a single LPC order and GD window size. To objectively choose their values we used (11). The rationale is that, given a good set of candidates, it is possible for the algorithm to select a good subset and λ would be low. Conversely, for a bad set no matter how efficient the algorithm is, λ would be high. In other words, the best value of λ from a set of candidates is dictated by the set itself.

$$Cost_x = \frac{1}{J}\sum_{j=1}^{J} 4\left(\frac{\tau_{ref_j} - \tau_{cl_j} + N_b}{l_j}\right)^2 \qquad (11)$$

where x denotes the set of reference CIs used, j is the stride number in a recording of J strides long, l_j is the jth stride length in samples, N_b is the bias in number of samples, τ_{ref_j} and τ_{cl_j} are the sample instances of the reference and its closest candidate, in stride j, respectively. We tested for N_b =-100 to 100 ms with 1 ms steps at 16 kHz sampling frequency for LPC orders 2 to 20 and GD window sizes 0.05 to 1 (as a fraction of the stride rate estimated using autocorrelation). For practical reasons the 0.05 limit was chosen to limit the candidates to a number that is manageable by the DP algorithm. The overall cost for a single combination is the mean of $Cost_{ici}$ and $Cost_{lci}$. The minimum is found at N_b =-5 ms which corresponds to order of 2 and window size of 0.05 (see Fig. 2). Sizes larger than 0.65 generate candidates that are not suitable for detecting both CIs (not shown). To validate our choice we ran the algorithm for sizes 0.03-0.07 and compared the results based on $S = \beta(A/100) + (1-\beta)[(100-\lambda)/100]$, where A is the average IDR and β is a weight factor set to 0.5. Depending on the application, β can be larger to favour high IDR or smaller to favour low λ. The best choice is the one that yields the highest score, $\hat{S} = (S_{ici} + S_{lci})/2$. For OA knees the algorithm performs better with a window size of 0.07 and for healthy with a size of 0.06. Overall, the best performance with a score of 0.9124 is achieved with a size of 0.06. Table 1 and Figs. 3 and 4 summarise the results for the selected order (2) and window size (0.06).

	OA ICI	**OA LCI**	**Healthy ICI**	**Healthy LCI**
IDR (%)	100	97.06	99.52	94.73
MR (%)	0	1.47	0.48	3.17
FAR (%)	0	1.47	0	2.10

Table 1. Average IDR, MR and FAR

5. DISCUSSION AND CONCLUSIONS

The results show that, in the database tested, the algorithm has detected on average 99.52% of ICIs and 94.73% of LCIs with an identification error of 13.0% and 16.0% respectively for the healthy recordings. For OA recordings an average of 100% of ICIs and 97.06% of LCIs have been detected with an error of 14.6% and 17.0%. An analysis has been performed on the choice of suitable LPC order and GD window size. Using stride detection and segmentation based on the DP algorithm, stride synchronous analysis can now be performed.

6. REFERENCES

[1] "Osteoarthritis in general practice: Data and perspectives," Tech. Rep., Arthritis Research UK, 2013.

[2] N.E. Lane and D. J. Walace, *All about osteoathritis: the definitive resource for arthritis patients and their families*, Oxford University Press, 2002.

[3] W. E. Blodgett, "Auscultation of the knee joint," *Boston Med. Surg.*, vol. 146, no. 3, pp. 63–66, 1902.

[4] E. Bircher, "Zur diagnose der meniscusluxation und des meniscusabrisses," *Zentralbl. Chir.*, vol. 40, pp. 1852–1857, 1913.

[5] S. Tavathia, R. M. Rangayyan, C. B. Frank, G. D. Bell, K. O. Ladly, and Y.T. Zhang, "Analysis of knee vibration signals using linear prediction," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 9, pp. 959–970, 1992.

[6] M. L. Chu, I. A. Gradisar, L. D. Zavodney, and G. F. Bowling, "Detection of knee joint diseases using acoustical pattern recognition technique," *J. Biomechan.*, vol. 9, pp. 111–114, 1976.

[7] M. L. Chu, I. A. Gradisar, M. R. Railey, and G. F. Bowling, "An electroacoustical technique for the detection of knee joint noise," *Med. Res. Eng.*, vol. 12, no. 1, pp. 18–20, 1976.

[8] M. L. Chu, I. A. Gradisar, L.D. Zavodney, and G. F. Bowling, "Computer aided acoustical correlation of papathologic cartilage generated noise," in *30th Ann. Conf. Eng. Med. Biol.*, 1977, p. 175.

[9] M. L. Chu, I. A. Gradisar, and R. Mostardi, "A noninvasive electroacoustical evaluation technique of cartilage damage in pathological knee joints," *Med. Biol. Eng. Comput.*, vol. 16, pp. 437–442, 1978.

[10] Y. T. Zhang and R. M. Rangayyan, "Adaptive cancellation of muscle contraction interference in vibroarthrographic signals," *IEEE Trans. Biomed. Eng.*, vol. 41, no. 2, pp. 181–191, 1994.

[11] Y. Shen, R. M. Rangayyan, G. D. Bell, C. B. Frank, Y. T. Zhang, and K. O. Ladly, "Localization of knee joint cartilage pathology by multichannel vibroarthrography," *Med. Eng. Phys.*, vol. 17, no. 8, pp. 583–594, 1995.

[12] S. Krishnan, R. M. Rangayyan, G. D. Bell, and C. B. Frank, "Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 6, pp. 773–783, 2000.

[13] K. Umapathy and S. Krishnan, "Modified local discriminant bases algorithm and its application in analysis of human knee joint vibration signals," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 517–523, 2006.

[14] B. Mascaro, J. Prior, L-K. Shark, J. Selfe, P. Cole, and J. Goodacre, "Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints," *Med. Eng. Phys.*, vol. 31, pp. 1013–1022, 2009.

[15] L-K. Shark, H. Chen, and J. Goodacre, "Discovering differences in acoustic emission between healthy and osteoarthritic knees using a four-phase model of sit-stand-sit movements," *Open Med. Inform.*, vol. 4, pp. 116–125, 2010.

[16] L.-K. Shark, J. Goodacre, and H. Chen, "Knee acoustic emission: A potential biomarker for quantitative assessment of joint ageing and degeneration," *Med. Eng. Phys.*, vol. 33, pp. 534–545, 2011.

[17] H. Chen, *Discovery of acoustic emission based biomarker for quantitative assessment of knee joint ageing and degeneration*, Ph.D. thesis, University of Central Lancashire, 2011.

[18] Mike Brookes, Patrick A. Naylor, and Jon Gudnason, "A quantitative assessment of group delay methods for identifying glottal closures in voiced speech," *IEEE Trans. Speech Audio Process.*, vol. 14, 2006.

[19] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, Apr. 1975.

[20] R. Bellman and S. Dreyfus, *Applied Dynamic Programming*, Princeton, N.J.: Princeton University Press, 1962.