An Approach for Automatic Identification of Fundamental and Additional Sounds from Cardiac Sounds Recordings

Amit Krishna Dwivedi, and Esther Rodriguez-Villegas

***Abstract*—This paper presents an approach for automatic segmentation of cardiac events from non-invasive sounds recordings, without the need of having an auxiliary signal reference. In addition, methods are proposed to subsequently differentiate cardiac events which correspond to normal cardiac cycles, from those which are due to abnormal activity of the heart. The detection of abnormal sounds is based on a model built with parameters which are obtained following feature extraction from those segments that were previously identified as normal fundamental heart sounds. The proposed algorithm achieved a sensitivity of 91.79% and 89.23% for the identification of normal fundamental, S_1 and S_2 sounds, and a true positive (*TP*) rate of 81.48% for abnormal additional sounds. These results were obtained using the PASCAL Classifying Heart Sounds challenge (CHSC) database.**

I. Introduction

CARDIAC sounds, caused by the mechanical vibrations of the heart or turbulent blood flow, can be recorded using an electronic stethoscope or a microphone, from the surface of the chest. The information can be presented graphically in what is known as a phonocardiogram (PCG). Fundamental heart sounds – first heart sounds (S_1) and second heart sounds (S_2) - are the ones mainly observed in the phonocardiogram of a healthy subject. S_1 is the result of the mechanical activities of the mitral and tricuspid valves; while S_2 results from the activities of the aortic and pulmonary valves. During normal cardiac conditions, S_1 appears as a single sound with discrete subcomponents M_1 and T_1; while S_2 appears as a single sound with discrete subcomponents A_2 and P_2, in the phonocardiogram. Although there are two discrete subcomponents in each of them (corresponding to the operation of the different valves), these are hard to distinguish because of the very small time interval existing between the occurrences of the individual mechanical cardiac events [1]. However, the time difference widens when there is malfunctioning of the heart valves [2].

Clinically, the splitting of S_1 sounds into its subcomponents, because of reasons other than the respiratory cycle, may be evidence of atrial spectral defects (ASD) or right bundle branch block (RBBB). Similarly, the splitting of S_2 sounds may be an indicator of left bundle branch block (LBBB), atrial spectral defects (ASD), right bundle branch blocks (RBBBs), left ventricular ectopic beats and pulmonary stenosis [3]–[5]. During cardiac abnormalities, apart from the split sounds, additional cardiac lub/dub sounds may be observed in the phonocardiogram, which can provide vital information about cardiac conditions and assist in the early stage detection of various cardiovascular diseases (CVDs) [5]. Consequently, the automatic detection of fundamental and additional heart sounds has been a topic of great interest among researchers [5]. However, as heart sounds are often overlapped with high-frequency acoustic signals unrelated to the operation of the heart, such as noise and artifacts, their correct interpretation is challenging. This also creates challenges when applying advanced signal processing methods to automate the process, especially, in real, non-controlled, environments where the number of disturbances present in the signal acquired is much higher.

This paper presents an approach to automatically segment cardiac peaks from sound recordings without the need of any additional ECG or pulse carotid signal; identify normal fundamental cardiac sounds; and differentiate them from both abnormal split sounds and additional lub/dub sounds in the systolic and diastolic intervals.

II. Segmentation of Heart Sounds Signals

A. Datasets

The results obtained in this paper were evaluated using sounds recordings obtained from the PASCAL Classifying Heart Sounds Challenge (CHSC) database [6]. 111 recordings marked with the exact locations of S_1 and S_2 sounds were used to validate the segmentation performance. Additionally, 19 recordings contained additional sounds either in the systolic or diastolic interval at regular intervals and 46 recordings contained additional sounds at irregular intervals. These were used to test the performance of the developed algorithm for the identification of additional sounds in the cardiac cycle. Apart from this dataset, heart sounds recorded in uncontrolled environment conditions using the digital stethoscope from Thinklabs [7], were used to validate the performance of the developed algorithm in closer to real life applications.

B. Segmentation of Heart Sounds Signals

The first step in the development of the algorithm was to localize peaks in the recorded signal. The temporal localization of all peaks would, later on, be used to assist in the identification of S_1 and S_2 sounds as well as other abnormal sounds. The following steps were followed:

This research was supported by the European Research Council (ERC), grant agreement no. 724334. The work of A. K. Dwivedi was supported by Imperial College London through the President's PhD Scholarship.

A. K. Dwivedi and E. Rodriguez-Villegas are with the Wearable Technologies Lab, Department of Electrical and Electronic Engineering, Imperial College London, SW7 2BT, United Kingdom. E-mail: {a.dwivedi16, e.rodriguez}@imperial.ac.uk.

1) Pre-processing: The amplitude of the heart sounds was normalized to a fixed scale of [-1, 1] dividing by its absolute maximum:

$$Y[n] = \frac{y'[n]}{\max_n(|y'[n]|)} \quad (1)$$

where, $Y[n]$ was the normalized signal and $\max_n(|y'[n]|)$ represented the absolute maximum of the resampled and filtered signal $y'[n]$. Further, a low pass filter with a cut off frequency of 500 Hz was employed, to eliminate high-frequency components from the normalized signals, which are typical of artifacts. The frequency spectrum of the recordings from the database was used to estimate this cut-off frequency (as shown in Fig. 1).

2) Identification of peaks of interest: As wavelets are rapidly decaying oscillations with zero mean, and well-suited for non-stationary signal analysis, wavelet analysis was chosen to obtain the time-frequency domain resolution of the cardiac sound signals. A 6th order Daubechies wavelet with 7-levels of decomposition and reconstruction was used. A windowing step was included to split the signal into a sequence of frames for the analysis. A 10 ms window was used to extract the envelope of the pre-processed signal using the average Shannon-energy (E_S):

$$E_S = \frac{-1}{N}\sum_{i=1}^{N} Y[n]^2(i).\,log Y[n]^2(i) \quad (2)$$

where, N represented the number of normalized samples in the selected window. The sound lobes of interest were localized using the zero crossing points of the normalized Shannon-energy E_{Snorm}(t), extracted as:

$$E_{Snorm}(t) = \frac{E_S(t) - \overline{E_S(t)}}{\sigma(E_S(t))} \quad (3)$$

where, E_S(t) was the average Shannon energy and $\overline{E_S(t)}$ and $\sigma(E_s(t))$ represented the mean and standard deviation of E_s(t), respectively.

III. Feature Extraction

Once the peaks were found, features were extracted to try to identify both, the presence of fundamental heart sounds and also additional heart sounds, corresponding to abnormalities.

A. Features to identify fundamental heart sounds

Evaluated features to extract the fundamental heart sounds (Fig. 2) included the time intervals between the adjacent peaks, ratio of the intervals, amplitude and frequency of each Shannon envelope, locations of the onsets and offsets of the Shannon envelope for each peak on the zero crossings, zero crossing rates, spectral width and the energy of the sound lobes.

B. Features to identify split sounds

The algorithm extracted the onset and offset measurements of each sound lobes of interest, which then were used to accurately determine the beginning and end of S_1 and S_2 sounds. Features to identify the split sounds were based on the consistency of S_1 and S_2 sounds. This was previously investigated by Tang *et al* [8]. Consecutive S_1 and S_2 sounds were aligned and plotted on top of one another to extract features. The onset and offset measurements obtained as the intersection of the Shannon envelope of each peak with the zero crossings were used to estimate the timing intervals of S_1 and S_2 sounds. This feature facilitated the identification of splitting sounds by evaluating the change in the intervals, with respect to the normal cardiac conditions, of the consecutive cardiac cycles.

Statistical features evaluated were the mean, standard deviation of skewness, mean absolute deviation, root mean square and root sum of squares of successive lobes and kurtosis from each analysis window. Apart from these statistical features, perceptual features i.e. mel-frequency cepstral coefficients (MFCCs) were extracted to capture the non-linear behavior of the signals.

IV. Identification of Normal Fundamental Heart Sounds

Sound lobes of interest in cardiac signals available from the database [6], as well as the signals acquired in uncontrolled environment, were identified using the wavelet transform and Shannon energy estimations. As an illustration, Fig. 3 shows the localization of sound lobes in an acoustic signal recorded using an electronic stethoscope [7] in a real, non-controlled environment.

As the duration of S_1 and S_2 sounds (without splitting of subcomponents) is known to range between 80 ms to 150 ms during normal cardiac operation, sound lobes with shorter or longer durations were investigated as potentially abnormal heart sounds [5]. The duration and energy of adjacent peaks were also evaluated, as S_1 and S_2 sounds cannot have the same energy under normal cardiac conditions [9]. Hence, if two adjacent peaks were found to have the same energy, then one of the two sounds was considered as abnormal. The intervals between the two adjacent S_1 and S_2 peaks represent the systolic and diastolic intervals, where the diastolic period under normal cardiac operation is greater than the systolic period. The start of S_1 sound marks the beginning of the

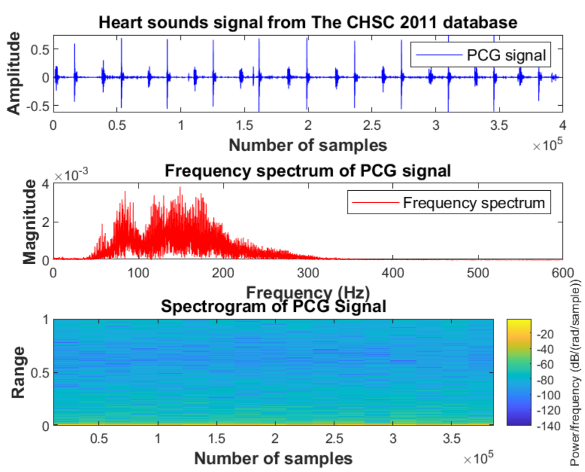

Fig. 1. Frequency spectrum using the Fourier analysis and the time-frequency analysis of heart sounds signal.

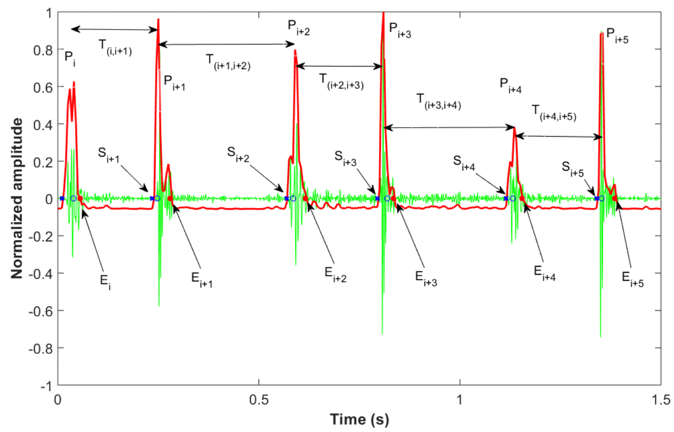

Fig. 2. Features extracted from heart sounds signals for the identification of S_1 and S_2 sounds. (S_i and E_i indicate the start and end locations of *i*th peak (P_i) and $T_{(i,i+1)}$ represents the time interval between peaks P_i and P_{i+1}).

systolic period and extends up to the beginning of S_2 sound; while the diastolic period is marked as the starting of S_2 sound and extends up to the starting of S_1 sound. The energy of sound lobes identified in the systole or diastole intervals regions was evaluated to distinguish the peak from the fundamental heart sounds. An example of how the transients were localized is shown in Fig. 4, in which the sound lobes of S_1 and S_2 were identified clearly in the signal obtained from the CHSC database [6]. Further, MFCCs were used to validate the sound lobes identified and to classify S_1 and S_2 sounds.

V. Approach to Identify Splitting of S_1 and S_2

Previous studies identified the discrete subcomponents of S_1 and S_2 using instantaneous amplitude and instantaneous phase information [3], perceptual features with k-means algorithm [10], linear and non-linear transient chirp models [1], damped sinusoid models [11], [12], and matching pursuit methods [13]. Following a chirp based method, the fundamental sounds are expressed mathematically as:

$$S_1(t) = A_{M_1}(t)\,sin\left(\varphi_{M_1}(t)\right) + A_{T_1}(t-\ \delta_{S1})\,sin\left(\varphi_{T_1}(t-\ \delta_{S1})\right) \quad (4)$$

$$S_2(t) = A_{A_2}(t)\sin\left(\varphi_{A_2}(t)\right) + A_{P_2}(t-\ \delta_{S2})\sin\left(\varphi_{P_2}(t-\ \delta_{S2})\right) \quad (5)$$

where, $\mathrm{A_{M_1}(t)\sin\left(\varphi_{M_1}(t)\right)}$ and $\mathrm{A_{T_1}(t-\ \delta_{S1})\sin\left(\varphi_{T_1}(t-\ \delta_{S1})\right)}$ represent the amplitude and phase of M_1 and T_1 components of S_1 sounds, respectively. Likewise, $\mathrm{A_{A_2}(t)\sin\left(\varphi_{A_2}(t)\right)}$ and $\mathrm{A_{P_2}(t-\ \delta_{S2})\sin\left(\varphi_{P_2}(t-\ \delta_{S2})\right)}$ represent the amplitude and phase of A_2 and P_2 components of S_2 sounds, respectively. Additionally, δ_{S1} and δ_{S2} are used to indicate the delay between the subcomponents of S_1 and S_2, respectively. Consecutive S_1 and S_2 sounds were aligned and plotted on top of one another (as shown in Fig. 5), to extract statistical information of sound lobes. These were extrapolated to obtain the characteristics of fundamental heart sounds in a cardiac signal.

As S_1 and S_2 sounds are quasi-cyclic stationary with almost fixed duration in consecutive cycles for a particular subject [8], a proper estimation of the spectral width of lobes helps in accessing the time intervals of S_1 and S_2 and the delay (δ_{S1} and δ_{S2}), which can be compared with the normal intervals (estimated during normal cardiac conditions) to ascertain the

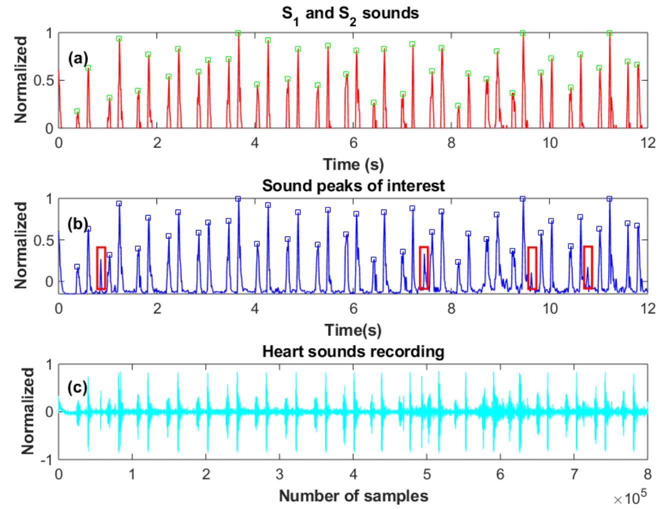

Fig. 3. Identification of S_1 and S_2 sounds in the heart sounds signals recorded in an uncontrolled environment with an electronic stethoscope. (c) and (b) show the recorded signal and the identified sound lobes of interest, respectively. (a) shows the potential S_1 and S_2 sounds after discarding the identified abnormal sounds.

splitting of heart sounds.

In case of splitting of S_1 or S_2, the interval between the two discrete consecutive sound lobes would be less than 50 ms and both sound lobes would have different energies [9]. This is because one of the split components of S_1 or S_2 will have a lower intensity compared to the another component [9]. Further, the consecutive sound lobes with time interval greater than 50 ms, were investigated for abnormal/additional lub/dub sounds. Since the possibility existed that split sounds of S_1 and S_2 might get confused with other abnormal heart sounds, perceptual features were also extracted using higher order decomposition of transients.

VI. Results

This proposed algorithm achieved an accuracy of 91.79% and 89.23% in successful identification of normal S_1 and S_2 sounds, respectively, which is superior to other reported envelope-based approaches [14]. Additionally, the proposed

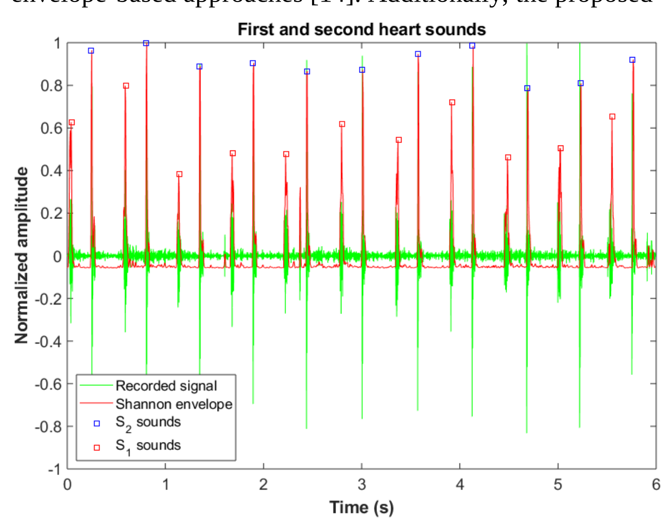

Fig. 4. S_1 and S_2 sounds identification using the proposed algorithm.

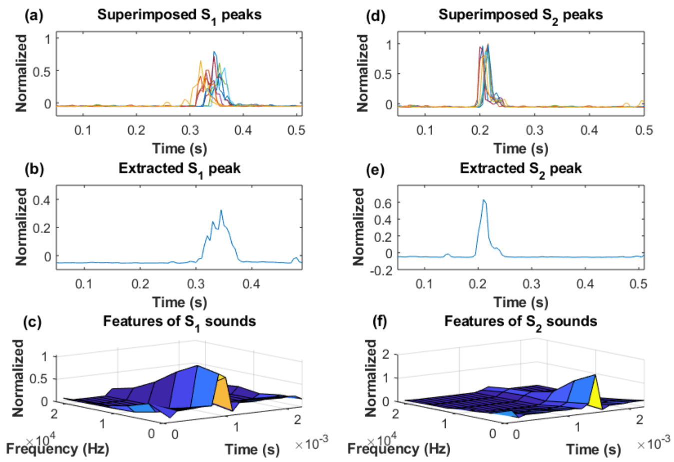

Fig. 5. S_1 and S_2 sounds aligned to extract features based on the consistency of S_1 and S_2 sounds. (a) and (d) show 22 S_1 and S_2 sounds from the database, aligned and plotted on top of one another to extract S_1 and S_2 sounds as in (b) and (e), respectively. (c) and (f) show the extracted features of S_1 and S_2 sounds, respectively.

algorithm is able to segment heart sounds signals without using any auxiliary signal, such as ECG and/or carotid pulse, which in other works is required to mark the location of S_1 and S_2 sounds.

Further, the performance of the proposed approach, using the extracted features was also tested with the recordings containing abnormal additional sounds, from the CHSC database [6]. The algorithm achieved a mean true positive (TP) rate of 81.48%, using an ensemble of k-NN classifier in the identification of additional lub/dub sounds in the systole or diastole interval with a 50-fold cross-validation. An example of how the transients were localized is shown in Fig. 6, in which sound lobes of S_1 and S_2 were identified clearly in the signal obtained from the CHSC database [6] that contained additional lub/dub sounds. A direct comparison of

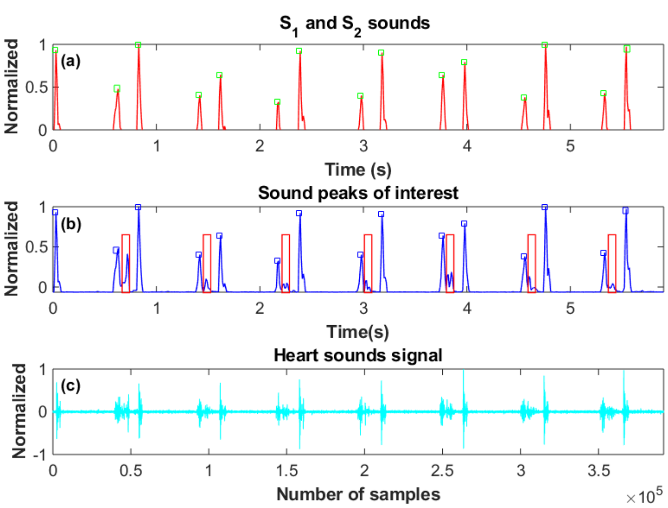

Fig. 6. Identification of S_1 and S_2 sounds in the heart sounds signals with additional lub/dub sounds obtained from the database. (c) and (b) show the recorded signal and the identified sound lobes of interest, respectively. (a) shows the potential S_1 and S_2 sounds after discarding the identified additional lub/dub sounds.

performance cannot be carried out since others have not tested their approaches with additional lub/dub sounds.

VII. Conclusions

This pilot study shows an approach to automatically segment heart sounds into cardiac events and from there differentiate between normal fundamental sounds and additional abnormal sounds; where the latter include both splitting of fundamental, S_1 and S_2, sounds and additional lub/dub sounds in the diastolic and systolic periods. The splitting of fundamental sounds is detected without extracting the discrete subcomponents of S_1 (M_1 and T_1) or S_2 (A_2 and P_2) sounds. This method could have a potential application in the automatic detection of cardiac abnormalities. It should be noted that, although, the proposed approach performed well on signals acquired in non-controlled environments, other features to assist with the extraction of subcomponents of S_1 and S_2 sounds could lead to further improvements.

References

[1] J. Xu, L. G. Durand, and P. Pibarot, “Nonlinear Transient Chirp Signal Modeling of the Aortic and Pulmonary Components of the Second Heart,” *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 1328–1335, 2000.

[2] S. M. Debbal and F. Bereksi-Reguig, “Analysis and study of the variation of splitting in the second heartbeat sound of wavelet transform,” *J. Med. Eng. Technol.*, vol. 30, no. 5, pp. 298–305, 2006.

[3] S. Barma, B. Chen, K. L. Man, and J. Wang, “Quantitative Measurement of Split of the Second Heart Sound (S2),” *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 12, no. 4, pp. 851–860, 2015.

[4] A. Djebbari and F. Bereksi-reguig, “Detection of the valvular split within the second heart sound using the reassigned smoothed pseudo Wigner–Ville distribution,” *Biomed Eng. Online*, vol. 12, no. 37, pp. 1–21, 2013.

[5] A. K. Dwivedi, S. A. Imtiaz, and E. Rodriguez-Villegas, “Algorithms for Automatic Analysis and Classification of Heart Sounds–A Systematic Review,” *IEEE Access*, vol. 7, pp. 8316–8345, 2019.

[6] P. Bentley, G. Nordehn, M. Coimbra, S. Mannor, and R. Getz, “The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011),” 2011. [Online]. Available: http://www.peterjbentley.com/heartchallenge/. [Accessed: 03-Mar-2017].

[7] “Thinklabs One Digital Stethoscope.” [Online]. Available: https://www.thinklabs.com/.

[8] H. Tang, T. Qiu, and T. Li, “Noise and Disturbance Reduction for Heart Sounds in Cycle-Frequency Domain Based on Nonlinear Time Scaling,” *IEEE Trans. Biomed. Eng.*, vol. 57, no. 2, pp. 325–333, 2010.

[9] S. Kang, R. Doroshow, J. McConnaughey, and R. Shekhar, “Automated identification of innocent Still’s murmur in children,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 6, pp. 1326–1334, 2017.

[10] T. Chen *et al.*, “S1 and S2 Heart Sound Recognition Using Deep Neural Networks,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 372–380, 2017.

[11] Y. Tang, C. Danmin, and L. G. Durand, “The Synthesis of the Aortic Valve Closure Sound of the Dog by the Mean Filter of Forward and Backward Predictor,” *IEEE Trans. Biomed. Eng.*, vol. 39, no. 1, pp. 1–8, 1992.

[12] A. Baykal, Y. Z. İder, and H. Köymen, “Distribution of Aortic Mechanical Prosthetic Valve Closure Sound Model Parameters on the Surface of the Chest,” *IEEE Trans. Biomed. Eng.*, vol. 42, no. 4, pp. 358–370, 1995.

[13] X. Zhang, L. Durand, S. Member, L. Senhadji, H. C. Lee, and J. Coatrieux, “Analysis-Synthesis of the Phonocardiogram Based on the Matching Pursuit Method,” *IEEE Trans. Biomed. Eng.*, vol. 45, no. 8, pp. 962–971, 1998.

[14] K. Hassani, K. Bajelani, M. Navidbakhsh, D. Doyle, and F. Taherian, “Heart sound segmentation based on homomorphic filtering,” *Perfusion*, vol. 29, no. 4, pp. 351–359, 2014.