

Algoritmi per l'identificazione automatica di eventi acustici anomali per l'Ambient Assisted Living

Emanuele Principi, Stefano Squartini, Francesco Piazza

Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche
Via Breccie Bianche 12, 60131, Ancona, Italy, e-mail s.squartini@univpm.it

Parole chiave: *Novelty Detection, Acoustic Monitoring, Machine Learning*

La possibilità di utilizzare sistemi automatici che supportino le persone nell'identificazione di eventi anomali, spesso associabili a situazioni di emergenza, sta riscontrando un notevole interesse dal punto di vista scientifico ed anche commerciale. In questa ricerca si punta l'attenzione sugli eventi acustici ambientali. L'obiettivo è di studiare algoritmi avanzati per l'identificazione automatica di quegli eventi acustici che sono legati a fenomeni che esulano dalle consuete attività quotidiane e che sono invece legati a situazioni di emergenza più o meno grave, come la caduta di una persona, la richiesta di aiuto o una intrusione. Esistono varie soluzioni in letteratura che vanno sotto il nome di *Novelty Detection* [1-3] e che vedono il coinvolgimento di algoritmi di Machine Learning. Alcune di esse sono orientate al settore della acoustic surveillance [4,5].

In questa ricerca gli autori, sfruttando le competenze assunte dalle ricerche degli anni passati in ambito speech processing, con particolare riferimento allo sviluppo di un framework distribuito per il riconoscimento di comandi e di richieste di aiuto [6-8], hanno affrontato il problema della novelty detection acustica per l'Ambient Assisted Living (AAL). Prendendo spunto dal lavoro di Ntalampiras et al. [5], è stato sviluppato un algoritmo che presenta le seguenti caratteristiche:

- Estrazione di *features* acustiche che permettano di comprimere il segnale d'ingresso mantenendone tutti gli aspetti informativi utili per l'identificazione dell'anomalia acustica;
- Creazione del modello di background acustico ("normality model" come riportato in Fig.1) sfruttando un approccio probabilistico (Gaussian Mixture Models - GMM o Hidden Markov Models - HMM) e per mezzo di un database opportuno (fase di training);
- Identificazione dell'occorrenza di novelty tramite il calcolo della "likelihood" statistica sulla base del modello di background generato (fase di testing – come in Fig.1); l'algoritmo include anche una strategia di adattamento del modello basato su approccio MAP (Maximum A-Posteriori).

Per la sperimentazione sono state considerate diverse *features*, ed in particolare: Mel-Frequency Cepstral Coefficients (MFCC), Power Normalized Cepstral Coefficients (PNCC – rappresentano una evoluzione rispetto alle precedenti e sono particolarmente indicate ad operare in presenza di rumore e riverbero [6]), Teager Energy Operator (TEO) e MPEG-7 *features* (Audio Waveform Type Descriptor, Audio Spectrum Flatness Descriptor, Audio Fundamental Frequency Descriptor).

Il database utilizzato è l'A3Novelty, sviluppato dagli autori di questa memoria e da altri validi collaboratori (<http://a3lab.dibet.univpm.it/research/a3novelty>). Il database è stato registrato in un laboratorio dell'Università Politecnica delle Marche, raccoglie più di 56 ore di registrazione in contesto diurno e notturno ed è caratterizzato da due tipologie di registrazioni audio: background e background con eventi sonori anomali (prodotti per mezzo di altoparlanti distribuiti nella stanza). Tali eventi sono raggruppati in 4 categorie: allarmi, cadute di oggetti, rotture di oggetti, urla.

Sono stati eseguiti numerosi test, coinvolgendo i due modelli statistici GMM e HMM per la caratterizzazione del background e le feature menzionate sopra. In Fig. 2 vengono riportate le Receiving Operating Characteristic (ROC) curve nel caso HMM (quello più performante), per i due contesti operativi (diurno/notturno), al variare delle *features* utilizzate. In fase di testing, scegliendo le sole PNCC in ingresso e adottando una strategia hard per la selezione della soglia decisionale, si è stato in grado di ottenere una True Detection Rate (TDR) pari al 93.75% ed una False Detection Rate (FDR) pari allo 0.59% nel caso diurno, e TDR=100%, FDR=5.37% nel caso notturno.

Sviluppi futuri sono orientati al coinvolgimento di altre *features* per una migliore caratterizzazione degli aspetti non-stazionari, all'ampliamento del dataset ad altri contesti AAL, alla introduzione di conoscenza a priori per particolari tipi di suono (come la voce) e allo sfruttamento di più canali microfonici per irrobustire le funzionalità dell'algorithm in scenari reali [9].

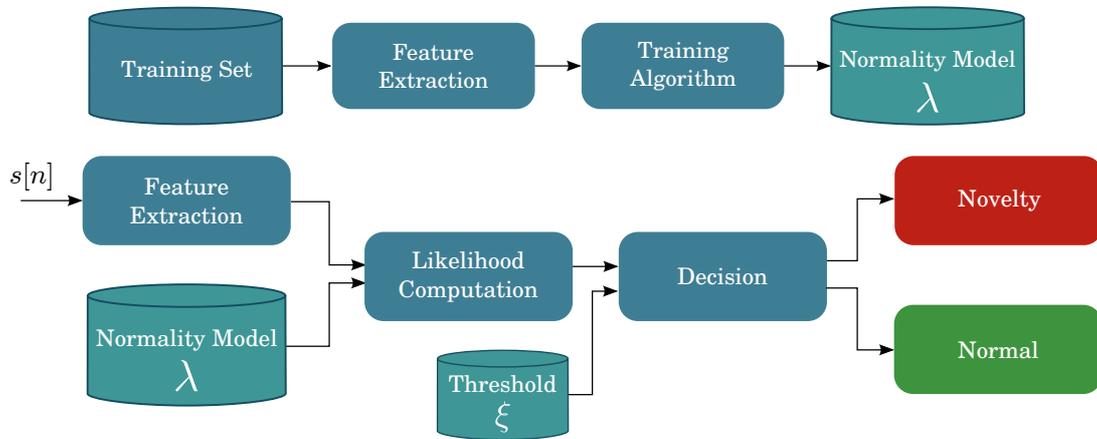


Fig. 1: Schema a Blocchi dell'algorithm di Novelty Detection acustica: fase di training (sopra) e di testing (sotto).

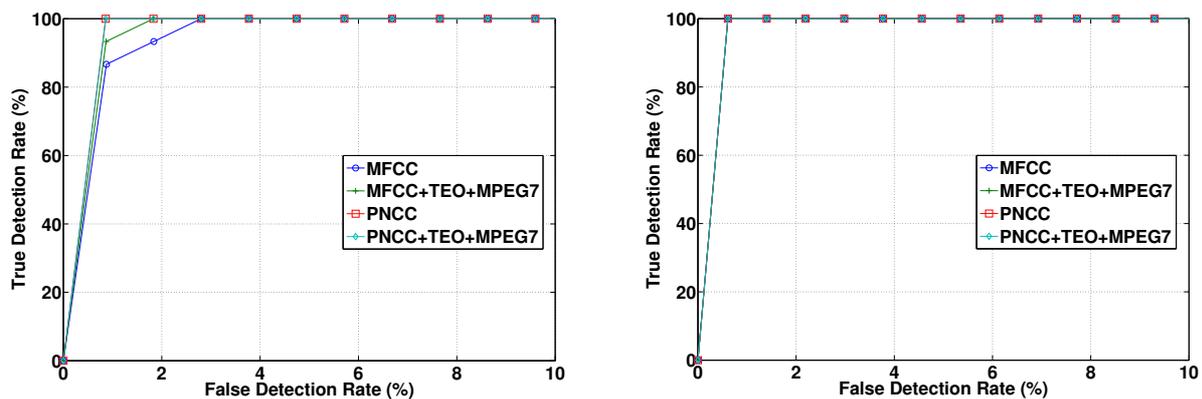


Fig. 2: ROC relative ai due case study considerati: diurno (sinistra) e notturno (destra).

BIBLIOGRAFIA

- [1] M. Markou and S. Singh. 2003a. Novelty detection: a review – part 1: statistical approaches. *Signal processing* 83, 12 (2003), 2481–2497.
- [2] M. Markou and S. Singh. 2003b. Novelty detection: a review – part 2: neural network based approaches. *Signal processing* 83, 12 (2003), 2499–2521.
- [3] M. A. F. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko. 2014. A Review of Novelty Detection. *Signal Processing* 99 (Jun. 2014), 215–249.
- [4] S. Ntalampiras, I. Potamitis, and N. Fakotakis. 2009. An adaptive framework for acoustic monitoring of potential hazards. *EURASIP Journal on Audio, Speech, and Music Processing* 2009 (2009), 13.
- [5] S. Ntalampiras, I. Potamitis, and N. Fakotakis. 2011. Probabilistic Novelty Detection for Acoustic Surveillance Under Real-World Conditions. *IEEE Trans. Multimed.* 13, 4 (Aug. 2011), 713–719.
- [6] E. Principi, S. Squartini, F. Piazza, D. Fuselli, and M. Bonifazi. 2013. A Distributed System for Recognizing Home Automation Commands and Distress Calls in the Italian Language. In *Proc. of Interspeech*. Lyon, France, 2049–2053.
- [7] E. Principi, D. Fuselli, S. Squartini, M. Bonifazi, and F. Piazza. 2013. A Speech-Based System for In-Home Emergency Detection and Remote Assistance. In *Proc. of the 134th International AES Convention*. Rome, Italy.
- [8] E. Principi, Roberto Bonfigli, S. Squartini, and F. Piazza. 2014. Improving the performance of a in-home acoustic monitoring system by integrating a vocal effort classification algorithm. In *Proc. of the 136th International AES Convention*. Berlin, Germany, to appear.
- [9] S. Squartini, E. Principi, R. Rotili, and F. Piazza. 2012. Environmental robust speech and speaker recognition through multi-channel histogram equalization. *Neurocomputing* 78, 1 (2012), 111–120.