

AGRICULTURAL ENGINEERING

1. La FIANZA FRANCESCO

Titolo tesi: Crop residue management: definition of internal standard for chopping quality and residue distribution evaluations for automation system development

Abstract

Con la necessità di nutrire una popolazione in crescita, gli agricoltori devono migliorare la produttività agricola senza compromettere la salute del suolo. Su questa base, la gestione sostenibile dei residui colturali sta diventando sempre più importante. I residui colturali, il materiale vegetale lasciato sul campo dopo il raccolto, svolgono un ruolo cruciale per la salute del suolo, il sequestro del carbonio e il controllo dell'erosione. La loro corretta gestione può migliorare la qualità del suolo, la produttività dell'azienda e l'impatto ambientale. Tuttavia, la sua gestione non è banale: una copertura eccessiva può impedire l'interazione suolo-seme, rendere l'azoto inaccessibile per l'assorbimento delle piante e mantenere il suolo freddo e umido; d'altro canto, una copertura insufficiente dei residui può degradare la qualità del suolo, contribuire alla perdita di materia organica e aumentare l'erosione. Quando si parla di gestione dei residui colturali, ci si riferisce a una strategia basata sulla riduzione della frequenza e dell'intensità delle operazioni di lavorazione del terreno e sull'aumento della quantità di residui dei raccolti precedenti. L'obiettivo di questo approccio gestionale è quello di preservare la qualità del suolo e dell'acqua, offrendo al contempo numerosi altri vantaggi ecologici e finanziari. Il passaggio alla gestione dei residui colturali può portare a rese più elevate e a un minore utilizzo di input come pesticidi, fertilizzanti, acqua di irrigazione e carburante, garantendo una gestione più redditizia del terreno. La gestione dei residui colturali ha l'obiettivo di migliorare o mantenere la fertilità e la salute generale del suolo a lungo termine. L'ecosistema del suolo è davvero complesso e il materiale lasciato sul campo dopo il raccolto o la lavorazione del terreno influenza diversi meccanismi e cicli naturali. Da un punto di vista fisico, può prevenire e mitigare l'erosione del suolo, mantenerlo a temperature favorevoli per la crescita delle piante, ridurre la compattazione e migliorare la sua capacità di trattenere e drenare l'acqua, mentre da un punto di vista chimico, preserva la fertilità del terreno arricchendolo di materia organica, riciclando i nutrienti e rendendoli disponibili per le piante, creando un habitat favorevole per i microrganismi. Va sottolineato che, se i benefici economici si notano quasi subito, quelli ambientali richiedono un periodo più lungo, ma sono anch'essi tangibili e forse ancora più importanti: la biodiversità e la conservazione degli habitat, insieme alla capacità di mitigare i cambiamenti climatici (siccità o forti piogge), sono oggi fattori estremamente importanti. Per massimizzare i benefici complessivi, gli agricoltori dovrebbero adottare un approccio olistico alla gestione dei residui che trovi un punto d'incontro tra fattori economici e ambientali. Le pratiche di gestione dei residui colturali comunemente utilizzate sono diverse e ognuna ha le sue peculiarità, rendendo il successo di una particolare strategia di gestione dipendente dal suolo, dal clima e dalle caratteristiche della pianta, tra gli altri fattori. Secondo alcuni agricoltori, quindi, un potenziale svantaggio della gestione dei residui colturali è che ciò che funziona in un ambiente può fallire in un altro. Per raccogliere i vari benefici della gestione dei residui colturali, gli agricoltori e gli specialisti agricoli devono condividere le loro competenze, fare le ricerche necessarie e utilizzare le moderne tecnologie. L'obiettivo del noi del dipartimento Harvesting Engineering di AGCO è quello di definire standard interni per la valutazione di parametri oggettivi relativi alla gestione dei residui colturali, per poter poi definire in modo inequivocabile i requisiti di prodotto su cui si baserà lo sviluppo di nuovi prodotti, abbracciando la strategia "farmer-first" di AGCO grazie alla quale lavoriamo insieme ai nostri clienti per comprendere a fondo le esigenze e allineare la nostra attività per soddisfare le loro aspettative e garantire un'agricoltura sostenibile per le generazioni future.

2. PACE GUIDO GIUSEPPE

Titolo tesi: Agrivoltaics in Italy: Technical Assessment and Simulation

Abstract

L'agrivoltaico è un nuovo modo di intendere l'energia rinnovabile, che unisce il fotovoltaico convenzionale con la contemporanea conduzione agricola, sotto e fra i pannelli fotovoltaici. Questa tecnologia sta assumendo un ruolo di primo piano nella transizione energetica italiana, in particolare per raggiungere i piani di decarbonizzazione fissati per il 2030. I campi agrivoltaici sono ancora pochi nel mondo e presentano caratteristiche costruttive diverse fra loro, poiché non si sono ancora affermati tipi dominanti di impianto né buone pratiche condivise. Per questo motivo, sono state indagate le potenzialità di una specifica configurazione di agrivoltaico, già presente nel mercato italiano, utilizzando come casi studio tre diversi impianti in fase di autorizzazione. Per prima cosa, è stata simulata la radiazione incidente al suolo, per comprendere la distribuzione della luce all'interno dell'impianto e individuare le zone più adatte alla crescita delle colture. Quindi, è stata indagata l'efficienza d'uso del suolo, per scoprire se l'utilizzo combinato dei terreni a fini agricoli e energetici possa essere preferibile rispetto a mantenere la separazione tra queste due attività: agricoltura da un lato e impianti fotovoltaici dall'altro. Infine, è stata svolta un'analisi economica per stimare i ritorni finanziari degli investimenti connessi agli impianti agrivoltaici, derivanti sia dalla componente energetica, sia da quella agricola. La simulazione della radiazione solare ha evidenziato una caratteristica forma a cupola nella distribuzione della luce al suolo, i cui valori massimi sono stati influenzati principalmente dalla distanza fra le file di moduli fotovoltaici. Sono emerse anche differenze relative alla localizzazione geografica degli impianti. Inoltre, tutti i casi studio hanno evidenziato un uso più efficiente del terreno nel caso di impianti agrivoltaici, con differenze nei risultati dovute principalmente al tipo di attività agricola. Infine, l'analisi economica ha mostrato ritorni dall'investimento in tempi più brevi in caso di agrivoltaico, quando confrontato con impianti fotovoltaici tradizionali, e i flussi di cassa finali sono risultati i più alti in due casi su tre.

3. REMELLI ANDREA

Titolo tesi: Technical assesement of Grana Padano cheese whey valorisation by purple phototrophic bacteria

4. VICENZONI RICCARDO

Titolo tesi: Uncovering the power of hyperspectral reflectance to predict plant community richness

Abstract

Questo lavoro di tesi mira ad analizzare le firme iper-spettrali di ecosistemi naturali terrestri, applicando indicatori esistenti in letteratura e di nuova concezione in modelli lineari multivariati per predire la ricchezza in specie vegetali. Il capitolo inerente lo stato dell'arte fornisce una revisione riguardo alcuni casi studio recenti, al fine di proporre differenti approcci di valutazione della variabilità spettrale di comunità vegetali, assieme alla spiegazione dei principi biologici sottostanti ai dati spettrali; specificamente, il capitolo esamina quattro articoli di ricerca che sono stati fondamentali nel plasmare il contributo sperimentale di questa tesi. Il capitolo sullo stato dell'arte si conclude con una caratterizzazione approfondita della tecnica di partizionamento della varianza spettrale, considerata uno dei metodi più affidabili in letteratura per analizzare consistentemente i gradienti alfa e beta di biodiversità. Il lavoro sperimentale della tesi è descritto nei capitoli materiale-e-metodi e

risultati-e-discussione, tramite l'esposizione di ogni fase del processo di analisi spettrale assieme alla descrizione dei tipi di dati raccolti e degli aspetti innovativi dell'analisi rispetto la letteratura di riferimento. L'originalità del suddetto lavoro risiede nell'aumentata complessità dei modelli multivariati, che incorporano variabili predittive di altezza della vegetazione con metriche di analisi spettrale delle medesime comunità vegetali: gli indicatori spettrali comprendono il partizionamento alfa della diversità spettrale tramite l'analisi di varianza, ed una nuova metrica di analisi entropica riguardo la distribuzione spaziale dei valori di riflettanza. Tale sperimentazione mira ad esplorare l'interazione tra metriche di altezza e spettrali in modelli lineari, al fine di identificare nuove percentuali di varianza nelle variabili risposta, dei suddetti modelli, che comprendono indicatori di ricchezza e di abbondanza delle specie vegetali.

5. ZAMBELLI RICCARDO

Titolo tesi: Field Simulator to Validate Automatic Cutting Height Control for Combine Harvesters

Abstract

Questa tesi presenta lo sviluppo e la convalida di un simulatore di controllo dell'altezza di taglio per mietitrebbie, condotto nell'ambito di uno stage presso AGCO S.p.A., nello stabilimento di Breganze, in Italia. Il simulatore è stato progettato per replicare le condizioni reali del campo, consentendo di testare e perfezionare gli algoritmi di controllo in un ambiente controllato. Il progetto mira a superare le sfide logistiche dei tradizionali test sul campo, spesso limitati dalla stagione del raccolto e dalle condizioni variabili del terreno. Lo sviluppo del simulatore ha comportato un approfondito studio preliminare dei sensori e dei segnali della mietitrebbia, assicurando che gli input simulati corrispondessero strettamente ai dati reali del campo. I componenti hardware e software sono stati sviluppati in parallelo, per integrare sensori laser ed inclinometro. Il simulatore è stato integrato con una mietitrebbia reale per registrare i profili del terreno e valutare le prestazioni del sistema di controllo in condizioni controllate e dinamiche. Inoltre, il simulatore ha permesso di generare manualmente segnali ideali, consentendo una valutazione precisa dei tempi di reazione del sistema. Le prestazioni del simulatore sono state convalidate utilizzando metodi di analisi statistica, tra cui la correlazione incrociata e la correlazione di Pearson, che hanno dimostrato un elevato grado di allineamento tra i segnali simulati e quelli reali. Il simulatore consente di utilizzare sia dati di prova reali sia segnali ideali creati manualmente, offrendo flessibilità nella valutazione delle prestazioni del sistema in condizioni di campo reale.

MUSIC AND ACOUSTIC ENGINEERING

1. ALFIERI RICCARDO

Titolo tesi: A Direct Data-Driven Approach for Active Noise Control of Tonal Disturbances

Abstract

Negli ultimi tempi, l'inquinamento acustico ambientale ha attirato sempre più attenzione, poiché è stato dimostrato che ha un impatto negativo sul benessere psicologico, mentale e fisico delle persone. Per affrontare questo problema, nel corso degli anni sono state sviluppate soluzioni di controllo del rumore sia passive che attive. Il controllo attivo del rumore (ANC) è efficace nelle bande di frequenza basse e medio-basse, dove le soluzioni passive falliscono a causa delle elevate lunghezze d'onda che le caratterizzano. L'algoritmo Filtered-x Least Mean Square (FxLMS) è il più comunemente utilizzato nei sistemi ANC. Questo algoritmo è classificabile tra gli approcci model-

based nel contesto di controllo del rumore attivo, poiché è necessario stimare la funzione di trasferimento del percorso secondario. Tuttavia, la stima di quest'ultimo non è sempre facile, soprattutto in sistemi tempo varianti. Per far fronte a questo problema, sono stati sviluppati algoritmi di controllo direct data driven, i quali sono indipendenti dal modello e possono quindi facilmente adattarsi ai cambiamenti del sistema. In questo lavoro si propone di sviluppare un sistema ANC basato su un algoritmo di controllo direct data driven: Simultaneous Perturbation Stochastic Approximation (SPSA). L'obiettivo è studiarlo e analizzarlo per comprendere se tali soluzioni possano essere utilizzate in casi reali. Il modello utilizzato in questo lavoro è una variante del tubo di Kundt a forma di Y. Inizialmente, il tubo di Kundt è stato modellato in COMSOL per estrarre i percorsi acustici e stimare le funzioni di trasferimento da utilizzare nelle simulazioni in Simulink. Successivamente, è stato sviluppato un nuovo metodo di ANC basato sull'algoritmo SPSA. Infine, è stata eseguita una validazione sperimentale confrontando il metodo proposto, basato sul controllo diretto, con l'algoritmo FxNLMS.

2. BESEDOVA OLGA

Titolo tesi: Multimodal emotion recognition system for piano performances

Abstract

I sistemi di riconoscimento delle emozioni stanno diventando sempre più importanti in vari settori, inclusi l'Interazione Uomo-Macchina (HMI), la musicoterapia e il miglioramento delle performance. Nel contesto musicale, riconoscere le emozioni trasmesse durante le esibizioni dal vivo al pianoforte offre preziose intuizioni sia sull'esperienza emotiva del musicista che dell'ascoltatore. Questa tesi affronta una lacuna nei modelli attuali di riconoscimento delle emozioni sviluppando un sistema che combina l'espressione musicale esterna con segnali fisiologici interni, al fine di catturare lo stato emotivo del performer durante esibizioni dal vivo al pianoforte. Integrando i segnali fisiologici, otteniamo una comprensione più approfondita delle emozioni in tempo reale del performer, offrendo una visione più chiara dei sentimenti alla base di una performance musicale. Un contributo chiave di questo lavoro è lo sviluppo di un framework multimodale per il riconoscimento delle emozioni, che combina i dati audio delle esibizioni dal vivo al pianoforte con segnali fisiologici, in particolare l'elettromiografia (EMG), per fornire un'analisi completa del contesto emotivo di una performance musicale. Il nostro modello utilizza una combinazione di tecniche di deep learning e fusione di dati multimodali per classificare gli stati emotivi con maggiore precisione. Per supportare e valutare le prestazioni del nostro sistema, introduciamo un nuovo protocollo per la costruzione di un dataset multimodale appositamente progettato per le esibizioni dal vivo al pianoforte, affrontando le limitazioni dei dataset esistenti. Questo dataset include audio sincronizzato, dati fisiologici del performer, registrazioni video focalizzate sul volto del performer e dati di movimento 3D, colmando una lacuna nelle risorse disponibili per questa ricerca. Attraverso una serie di esperimenti, dimostriamo che l'integrazione di segnali fisiologici migliora significativamente la precisione della classificazione delle emozioni rispetto ai sistemi basati solo sull'audio. Questo approccio mostra un grande potenziale per applicazioni come il miglioramento delle performance e le visualizzazioni musicali basate sul contesto emotivo in tempo reale. Questo lavoro rappresenta un contributo significativo al campo del riconoscimento delle emozioni, in particolare nel riconoscimento delle emozioni musicali, offrendo una nuova prospettiva su come le emozioni possano essere comprese combinando espressioni musicali con segnali fisiologici, aprendo la strada a future ricerche e applicazioni sia in ambito musicale che nelle aree legate alle emozioni.

3. COLOTTI FRANCESCO

Titolo tesi: Towards Joint Music Source Separation and Deep Drums Demixing

Abstract

Gli algoritmi di Music Source Separation (MSS) sono ampiamente usati, permettendo sia a musicisti che ricercatori di separare tracce mixate in stem, tipicamente quattro, fra cui voci, basso, batteria e altri. Il Drums Demixing (DDX) è un sottocampo di MSS che punta ad ottenere stem isolati delle singole parti della batteria da una traccia di sola batteria. Ad oggi, per estrarre sia le parti strumentali che di batteria da una canzone, è necessario applicare prima un algoritmo di MSS e, successivamente, un modello di DDX allo stem di batteria isolato dal primo. L'obiettivo di questa tesi è valutare la performance di un singolo modello che svolge contemporaneamente MSS e DDX, confrontandolo con gli attuali approcci a due stadi. Per far ciò, due architetture allo stato dell'arte sono state addestrate e testate come modelli di MSS (a 4 stem), DDX (a 5 stem) e MSS+DDX (a 8 stem). I risultati mostrano che le configurazioni a singolo stadio sono meno efficaci rispetto alle loro controparti a due stadi in termini di rapporto segnale/distorsione, richiedendo tuttavia circa metà del tempo per l'inferenza a parità di hardware, evidenziando un compromesso fra performance e requisiti di sistema.

4. CRISTOFALO MARCO

Titolo tesi: A model based active noise control logic exploiting virtual sensing and future signal prediction for reducing fast transient disturbances

Abstract

I sistemi di Riscaldamento, Ventilazione e Aria Condizionata (HVAC) sono fonti significative di rumore in ambienti confinati come abitacoli di veicoli e spazi industriali. Questo rumore, che si manifesta come ronzio a bassa frequenza, scoppi impulsivi o flussi d'aria turbolenti, può compromettere il comfort e la produttività. I tradizionali sistemi di Controllo Attivo del Rumore (ANC), sebbene efficaci in alcuni casi, incontrano difficoltà nell'attenuazione di rumori non stazionari e segnali caratterizzati da transienti veloci a causa del tempo di convergenza richiesto dagli algoritmi adattivi, che necessitano di diversi campioni del segnale di disturbo per calcolare un filtro efficace. Per affrontare queste limitazioni, la presente tesi introduce un algoritmo ANC che incorpora tecniche innovative, tra cui un metodo di virtual sensing basato su un filtro di Kalman, per stimare il segnale di disturbo in una determinata ubicazione spaziale dove non è installato alcun sensore, permettendo la cancellazione del rumore in tale posizione, a condizione che il modello sia accuratamente analizzato e modellato. Inoltre è eseguita la previsione futura del segnale di disturbo utilizzando uno strumento avanzato, noto come predittore dei sistemi a tempo morto, basato anch'esso sulla metodologia del filtro di Kalman. Questa previsione consente al sistema di adattare preventivamente il segnale di controllo anticipando possibili cambiamenti improvvisi. Il segnale di controllo è calcolato attraverso l'inversione della funzione di trasferimento del percorso secondario. Questo approccio garantisce una cancellazione del rumore precisa e tempestiva, contrastando le specifiche caratteristiche del disturbo. L'algoritmo proposto è valutato attraverso simulazioni numeriche e validazione sperimentale su un setup costituito da un tubo di Kundt in configurazione ad Υ . I risultati dimostrano significativi miglioramenti nell'attenuazione del rumore HVAC non stazionario, sottolineando l'efficacia dell'algoritmo contro segnali con transienti veloci e di tipo impulsivo. Le sue capacità suggeriscono potenziali applicazioni in vari contesti automobilistici e industriali.

5. GAFENCU SEBASTIAN CRISTIAN

Titolo tesi: Modeling Circuits with Multiple N-port Nonlinearities in the Wave Digital Domain

Abstract

Questa tesi presenta un metodo innovativo per la modellazione Virtual Analog (VA), un'area di ricerca che mira a emulare digitalmente apparecchiature audio analogiche. L'approccio proposto si concentra sui filtri digitali a onde (Wave Digital Filters, WDF) e riguarda una tecnica per gestire in modo efficiente elementi non lineari multi-porta multipli. Raggruppando le non-linearità in un unico blocco attraverso la definizione vettoriale delle onde, abbiamo dimostrato che è possibile modellare esplicitamente tali elementi nel dominio delle onde digitali utilizzando reti neurali semplici come i perceptron multistrato (Multi-Layer Perceptrons, MLP). In questo modo, il metodo evita la complessità computazionale introdotta dai solutori iterativi, che tipicamente rallentano la simulazione. Testiamo l'approccio proposto per l'emulazione di due celebri circuiti di effetti per chitarra, ovvero il Vox V847 Wah-Wah e l'Arbiter Electronics Fuzz Face. I risultati dimostrano che il metodo proposto non solo è caratterizzato da un'elevata accuratezza, ma offre anche miglioramenti significativi in termini di prestazioni in tempo reale, fornendo una soluzione generale, scalabile ed efficiente per la modellazione VA di circuiti con molteplici non-linearità multi-porta.

6. GIONFRIDDO MATTEO

Titolo tesi: Controllable music generation with neural discrete representations and a multitrack MIDI-to-audio dataset

7. LISTA DAVIDE

Titolo tesi: A multimodal approach to Artist Similarity assessment using Audio and Social Media content

Abstract

In un'epoca in cui l'industria musicale è sempre più influenzata dalla presenza online degli artisti, questa tesi propone un approccio innovativo per l'analisi della similarità tra artisti, combinando dati multimodali derivati dall'audio delle canzoni e dalle immagini e dai testi pubblicati su Instagram. Gli embeddings audio sono stati estratti utilizzando modelli di deep learning avanzati, come il modello VGGish di Google, applicato a tracce musicali pre-processate sotto forma di log-mel spettrogrammi. Per l'analisi dei contenuti social media, sono stati utilizzati il modello CLIP di OpenAI per le immagini e il modello SBERT per i testi delle didascalie dei post. Il fulcro della ricerca è una rete neurale siamese addestrata con triplet loss e cosine distance. Dopo aver fuso i dati multimodali raccolti in una rappresentazione unificata, questi sono stati elaborati da una rete siamese 1D CNN per calcolare una misura di similarità. La rete è in grado di mappare gli artisti in uno spazio latente di 256 dimensioni, collocando vicini gli artisti con caratteristiche simili, sfruttando sia le somiglianze identificate manualmente da esperti nel settore, estratte dalla piattaforma AllMusic, sia quelle algoritmiche estratte da Spotify, tramite le loro API, per la creazione di un dataset ampio e il più completo possibile. Gli esperimenti condotti su un campione di 3.008 artisti presenti su Instagram hanno dimostrato l'efficacia del modello multimodale. Le prestazioni del modello sono state valutate attraverso metriche come precision (o validità), recall e visualizzazioni t-SNE, mostrando che l'integrazione di dati visivi e testuali con i dati audio fornisce un contributo significativo alla catalogazione degli artisti rispetto ai metodi tradizionali basati solo sull'audio. I risultati indicano che un approccio multimodale

offre maggiore accuratezza e una separazione più netta tra artisti non simili, suggerendo un potenziale miglioramento nei sistemi di raccomandazione musicale e nella scoperta di nuovi artisti, e colmando il divario tra produzione musicale e presenza online.

8. MALAMAN VITTORIA

Titolo tesi: Analysis of a Locally Resonant Metamaterial for interior noise mitigation in high-speed trains

9. MASSIMI MATTIA

Titolo tesi: Physical Modeling of Timpani Using Digital Waveguide Meshes

Abstract

Questo studio presenta un approccio a modellazione fisica per ricreare il suono di un timpano utilizzando mesh di digital waveguides (DWG). Il metodo delle DWG, noto per la sua efficienza nella simulazione della propagazione delle onde negli strumenti musicali, è stato scelto per la sua capacità di riprodurre suoni realistici. Il modello simula accuratamente le relazioni armoniche e il comportamento dinamico di un vero timpano, offrendo grande versatilità grazie alla possibilità di regolare vari parametri fisici come la dimensione della membrana, il punto di percussione e il tipo di eccitazione. Integrando modelli e concetti da numerosi studi presenti in letteratura, è stato sviluppato un modello che risulta essere sia relativamente semplice che accurato. Inoltre, vengono esplorate diverse caratteristiche per migliorare il realismo del suono generato, ottenendo un modello finale che corrisponde strettamente agli armonici e al decadimento temporale di un vero tamburo. Le scelte progettuali, comprese varie semplificazioni, sono bilanciate tra la necessità di accuratezza sonora e i vincoli di potenza di calcolo e frequenze di campionamento. L'efficienza computazionale della mesh DWG a variabili K (KDWM), consente un'ottimizzazione significativa, inclusi aggiornamenti paralleli tra i nodi della mesh. Viene fornito anche un frammento di pseudocodice tratto dall'implementazione in MATLAB. Questo lavoro mira inoltre unire l'acustica teorica e la sintesi sonora pratica.

10. PETTENO' MATTEO

Titolo tesi: Latent Space Regularization via Normalizing Attribute Transformations for Symbolic Music Generation

Abstract

Nonostante l'innegabile successo di recenti modelli generativi come GPT-4 e DALL-E, essendo basati su prompt di testo, essi faticano a fornire un controllo preciso su attributi specifici del contenuto generato. Tale controllo tuttavia, è spesso auspicabile in applicazioni come la generazione di musica e la composizione musicale assistita. A questo scopo un approccio spesso seguito consiste nell'utilizzo di modelli a variabili latenti, i quali permettono di imparare nel loro spazio latente una particolare rappresentazione che evidenzia i diversi fattori di variazione nei dati di input, consentendo così una manipolazione più precisa delle proprietà dell'output durante la sua generazione. In questo lavoro estendiamo il framework conosciuto come variational information bottleneck con un approccio di tipo supervised: uno specifico attributo viene codificato in una dimensione dello spazio latente grazie ad un termine di regolarizzazione aggiuntivo nella funzione di loss globale. La novità della nostra proposta risiede nell'introduzione di una funzione parametrica e invertibile il cui obiettivo è quello di trasformare

la distribuzione originale e possibilmente complessa dell'attributo in oggetto in una più semplice che sia coerente con quella scelta per lo spazio latente. In particolare, tale trasformazione è progettata in modo da incoraggiare la sinergia tra il termine di regolarizzazione sopra citato e la divergenza di Kullback-Leibler durante la fase di training del modello. Inoltre, essendo invertibile, la trasformazione proposta garantisce maggior interpretabilità nel processo di manipolazione dell'attributo. Considerando diversi attributi, applichiamo il metodo proposto nel dominio della symbolic music generation, dimostrando che, rispetto allo stato dell'arte, conduce ad effettivi miglioramenti nella correlazione tra l'attributo delle melodie generate e la dimensione dello spazio latente in cui è codificato, preservando al contempo la distribuzione delle variabili latenti.

11. POPOVIC NATASA

Titolo tesi: Is JPEG-AI going to change multimedia forensics?

Abstract

Bilanciare il compromesso tra alta qualità delle immagini e dimensioni ridotte dei file è sempre stata una sfida significativa nel campo dell'imaging digitale. La compressione lossy JPEG ha costituito lo standard di riferimento per decenni, offrendo un compromesso ragionevole tra i due. Tuttavia, la necessità di soluzioni avanzate è diventata evidente, poiché dobbiamo memorizzare immagini più che mai, in una qualità superiore a quella di un tempo. La soluzione risiede nello sviluppo di tecniche di compressione delle immagini basate su reti neurali. Ri- conoscendo il potenziale e la necessità di questi progressi, il comitato JPEG ha deciso di sviluppare un nuovo standard per la compressione delle immagini basato sull'intelligenza artificiale (IA), noto come JPEG-AI, la cui pubblicazione e adozione è prevista per la fine del 2024. Ma, come tutti i nuovi sviluppi che portano grandi benefici, vengono sempre accompagnati da qualche tipo di problema. Le tracce e gli artefatti JPEG sono sempre stati elementi chiave nella forensica delle immagini come modo per determinare se un'immagine è originale o manipolata, ma la compressione basata sull'IA non opera secondo i metodi classici del JPEG. La maggior parte delle tecniche di compressione neurale si basa su autoencoder che comprimono le immagini codificandole in uno spazio latente e successivamente ricostruendole tramite un decodificatore, quindi gli artefatti introdotti sono intrinsecamente diversi. Questo studio mira a esaminare gli artefatti introdotti da JPEG-AI, includendo altre tecniche di compressione neurale come HiFiC, e a scoprire se rappresentano una sfida per la comunità forense multimediale. In particolare, dato che la letteratura ha dimostrato che gli artefatti generati da tecniche di compressione neurale presentano similarità con quelli rintracciabili nelle immagini generate da reti neurali, un aspetto cruciale è esplorare come JPEG-AI potrebbe fuorviare i detector forensi esistenti, e.g., confondendo le immagini neurali compresse reali con quelle sintetiche, o mascherando potenzialmente la manipolazione delle immagini come lo splicing. I nostri risultati confermano queste ipotesi e mirano a sensibilizzare i ricercatori forensi riguardo la futura adozione di questo standard.

12. RAPISARDA CLAUDIO

Titolo tesi: Audio Language Models for Deep Packet Loss Concealment

Abstract

Negli ultimi anni, la diffusione capillare delle tecnologie di streaming e VoIP ha determinato un aumento sostanziale dell'interesse sul tema del Packet Loss Concealment (PLC). L'impiego sempre più frequente dei cosiddetti Audio Language Models (ALM) in task di generazione automatica

condizionata ad un prompt di testo, inoltre, offre lo spunto per l'argomento su cui verte la nostra ricerca: proporre un approccio inedito al PLC basato proprio sugli ALM. Gli ALM offrono le stesse possibilità di parallelizzazione e la stessa ottima capacità di modellizzare sequenze dei Large Language Models (LLM) da cui derivano. Le loro capacità in ambito generativo sono ben documentate in letteratura, ma la loro applicabilità a task di next-token prediction, come è il caso di un sistema di PLC, è ciò che approfondiamo in questo nostro studio. Oltre a proporre un nuovo approccio, durante la nostra ricerca abbiamo ideato diversi pattern di interleaving, allo scopo di affinare la qualità predittiva del nostro modello e migliorarne le prestazioni. Ogni pattern è stato implementato all'interno di due ALM, un modello leggero, con meno di 10M di parametri, e un secondo più pesante, con circa 30M di parametri. Per ognuno di questi modelli abbiamo poi misurato l'accuratezza e la qualità del segnale predetto, oltre che il tempo di inferenza medio. I risultati dei nostri test mostrano come anche ALM leggeri siano efficaci nel contesto del PLC. Lo studio dimostra, inoltre, l'impatto notevole che i pattern di interleaving possono avere sui modelli, sia in termini di accuratezza che di latenza. Con questa ricerca speriamo di contribuire all'avanzamento delle tecniche di PLC e allo sviluppo di nuove applicazioni per gli ALM.

13. SGOTTO SILVIO

Titolo tesi: Virtual Acoustics and Auditory Perception Evaluation Applied to the Pre-Romanesque Church of Sant Miquel in Terrassa

Abstract

In un mondo in cui le tecnologie legate alla realtà virtuale sono di crescente interesse scientifico e culturale, la ricreazione virtuale di ambienti acustici reali ha guadagnato un ruolo cruciale nel donare all'ascoltatore un'esperienza immersiva. I più noti ed utilizzati software di acustica architettonica (ad esempio RAVEN, EASE ed ODEON) utilizzano algoritmi di ray tracing e metodi delle sorgenti immagine per simulare sia le prime riflessioni sia la riverberazione di ambienti virtuali complessi (importati nei software come modelli 3D). Questi software permettono così, in linea di principio, di simulare la risposta all'impulso (IR) di qualsivoglia ambiente reale, permettendo così la produzione di auralizzazioni, ovvero i risultati udibili di informazioni fisiche (ad esempio i parametri acustici dell'ambiente in questione). Nel presente lavoro di tesi si sono prodotte auralizzazioni della chiesa di Sant Miquel in Terrassa (vicino a Barcellona), la quale è stata scelta come esempio di ambiente in cui, in epoca pre-romanica (precedente quindi al secolo XI d.C.), si praticava il cosiddetto Rito Hispanico. A tal scopo un preesistente modello virtuale della chiesa è stato calibrato e utilizzato per simulare la IR della chiesa e applicarla così ad un set di audio anecoici. I risultati sono stati valutati in relazione ad analoghi audio registrati (con la stessa combinazione sorgente-ricevitore) nella chiesa reale e con altrettanti audio auralizzati tramite l'utilizzo della risposta all'impulso misurata nella chiesa. Il metodo di valutazione è stato sia fisico (basato su uno studio sullo spettro degli audio) sia percettivo. Per svolgere la valutazione percettiva si è ideata e svolta una campagna di test d'ascolto alla camera anecoica del Politecnico di Madrid - Campus Sur e alla camera acusticamente trattata del Politecnico di Milano - Campus di Cremona. I risultati sono stati analizzati avvalendosi della Teoria del rilevamento del segnale (SDT), e in particolare tramite la misura della differenza sensoriale d-prime tra audio registrato nell'ambiente reale, quello ottenuto tramite IR misurata e quello ottenuto tramite IR simulata.

14. VIVIANI MARCO

Titolo tesi: Hybrid Packet Loss Concealment for Speech Applications Based on Linear Predictive Coding and Deep Residual Learning

Abstract

Con \Packet Loss Concealment\ (PLC) ci si riferisce a qualsiasi tecnica progettata per correggere e mascherare i messaggi danneggiati nelle comunicazioni audio attraverso protocolli Internet causati da errori di trasmissione dei dati. Al giorno d'oggi la telefonia via IP e i sistemi di teleconferenza sono essenziali sia nella vita professionale che personale, ma sono tuttavia soggetti a perdite di dati e latenza. Da qui la necessità di tecniche PLC efficaci in grado di operare in tempo reale e fornire dati audio ad alta qualità. In questo lavoro presentiamo una tecnica ibrida in cui una rete neurale e un modello autoregressivo cooperano per ricostruire i pacchetti persi. Il modello in esame è PARCnet, una recente tecnica PLC che ha dimostrato ottimi risultati nel campo delle reti musicali in termini di percezione uditiva e inferenza in tempo reale. L'obiettivo principale di questo lavoro è quello di passare dall'ambito musicale a quello vocale, cercando di replicare gli ottimi risultati del modello originale grazie ad una nuova architettura e implementazione dei suoi due rami e anche a nuove strategie di training. Il nostro nuovo modello è stato testato con metriche oggettive allo stato dell'arte e con un test percettivo su un campione di ascoltatori. I risultati confermano che tale metodo supera le prestazioni delle baselines prese in considerazione e traccia un nuovo corso per affrontare il problema del PLC.

15. ZEZZA FEDERICA

Titolo tesi: Video Deepfake Detection through Head Pose Estimation

16. ZHUANG YAN

Titolo tesi: Can Blind AI Fight? Exploring AI and Audio Design in DareFightingICE

Abstract

L'obiettivo di questa tesi è presentare Blaze, un agente di deep reinforcement learning, insieme a due modelli di audio design denominati Adaptive ed Enhanced, sviluppati per le competizioni DareFightingICE all'IEEE CoG 2024. A differenza degli agenti all'avanguardia, che si basano principalmente su informazioni visive, Blaze utilizza esclusivamente input audio per il processo decisionale, sfruttando quattro sound encoder distinti. L'addestramento è stato effettuato utilizzando l'algoritmo Rainbow, permettendo a Blaze di apprendere e prendere decisioni strategiche. I risultati empirici mostrano prestazioni eccezionali, dimostrando la capacità dell'agente di comprendere lo stato del gioco attraverso segnali audio e di superare gli avversari, raggiungendo il secondo posto nella competizione AI DareFightingICE del 2024. Per esplorare la sound accessibility e migliorare il sound design nei giochi di combattimento per giocatori ipovedenti, proponiamo due approcci distinti. Il primo, Adaptive, è un sistema che regola il volume della musica di sottofondo in base a parametri di gioco come i punti salute e i punti energia. Il secondo, Enhanced, adotta un approccio più tradizionale, concentrandosi sugli effetti sonori migliorati. Abbiamo valutato l'efficacia di entrambi i design attraverso esperimenti condotti con persone e AI, utilizzando Blaze come uno dei giudici. I risultati hanno mostrato che il design Enhanced ha superato il vincitore della DareFightingICE sound competition del 2023, classificandosi terzo nell'edizione del 2024. Nel confronto tra i due approcci, i giocatori umani hanno ottenuto prestazioni migliori con il design Adaptive, mentre le AI hanno

preferito l'approccio Enhanced. Questi contributi alla ricerca evidenziano il potenziale delle soluzioni audio-based per migliorare l'accessibilità dei giochi per giocatori ipovedenti.