



FreeTW
il Tele-Web-Visore

Il riconoscimento vocale



Il riconoscimento vocale

Per usare i comandi vocali senza litigare con il computer se ne devono comprendere i limiti e accettarli.

- 1) Per far funzionare i comandi vocali su WindowsXP e' necessario fare una installazione aggiuntiva : il "Microsoft Speech SDK 5.1.msi" e questo richiede alcuni minuti di fastidio aggiuntivi.
- 2) Sui sistemi operativi "Vista" venduti per il mercato italiano i comandi vocali non ci sono e sembra che non esista nessun modo di installare qualcosa per farli funzionare. Che i "Vista" non funzionino in italiano e' comprensibile ma dovrebbero poter andare almeno in inglese !
- 3) In ogni caso i comandi vocali funzionano solo in inglese e non c'è da sperare che Microsoft si decida a fare la versione italiana perché sono dieci anni che lo promette e non li abbiamo ancora...
- 4) Per evitare al massimo la possibilità di falsi riconoscimenti le parole valide sono solo una ventina per cui ci si scordi di poter usare frasi fantasiose tipo "chiudi questo film e fammi vedere le foto.
- 5) E' necessario tenere il microfono abbastanza vicino e parlare abbastanza forte e non e' possibile alzare il volume dei film (o dalla musica) in modo esagerato, più il film e' alto e più si deve tenere il microfono vicino alla bocca e parlare forte.
- 6) Si deve usare un buon microfono, meglio se preamplificato, meglio ancora se preamplificato e direttivo.
- 7) Si deve controllare che il microfono funzioni bene e si devono controllare e regolare moltissime cose nel "Pannello di controllo" / "Sintesi e riconoscimento vocale"
- 8) Con i comandi vocali si possono vedere bene i film e le foto, ascoltare gli mp3 e le radio ma per navigare su internet e' meglio usare il telecomando.

Qualcuno troverà che i comandi vocali vanno bene e che sono "il massimo" mentre qualcun altro si lamenterà che non funzionano del tutto.

Il mio consiglio per chi si trova male con i comandi vocali e' di mantenerli disabilitati e di usare il telecomando che invece va bene in ogni situazione.

Regolare le proprietà del riconoscimento vocale

Le istruzioni seguenti presuppongono che siate interessati a parlare abbastanza piano ed a tenere il microfono abbastanza lontano in modo che non dia fastidio (venti o trenta centimetri)

Chi invece urla come un tricheco e/o tiene il microfono incollato alla bocca potrebbe generare un segnale che supera il massimo e dovrà regolarsi in modo diverso per evitare distorsioni.

Probabilmente si possono usare i comandi vocali senza fare niente altro che abilitarli non pannello opzioni di FreeTW ma se si vuole farli funzionare al meglio e' bene seguire con cura tutti i consigli seguenti.

Creare un nuovo utente

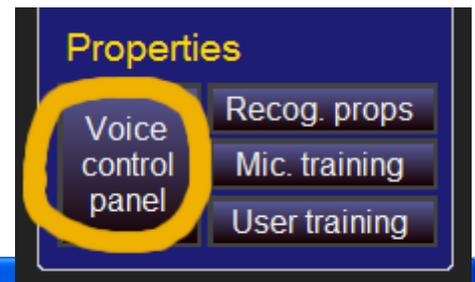
Se abbiamo appena installato il riconoscimento vocale ("Microsoft Speech SDK 5.1.msi") allora dovrebbe esistere un solo utente e questo utente dovrebbe essere vergine, cioè senza nessun apprendimento.

A noi serve un utente che sia sicuramente senza nessun apprendimento perché così' gli facciamo imparare le venti parole che ci servono, e solo quelle, e così' funziona molto meglio.

In ogni caso creare un nuovo utente e' un attimo, se non siamo del tutto sicuri di quelli che già' esistono e' sempre meglio fare un utente nuovo e poi cancellare tutti gli altri.

Questa procedura può anche servire per creare più di un utente, magari uno maschile e uno femminile, e poi sceglierli a seconda di chi parla.

Prima di tutto aprire il pannello "Opzioni" di FreeTW e premere il tasto "Voice control panel" che ci fa accedere alle proprietà del riconoscimento vocale (che sono le stesse accessibili dal pannello di controllo)



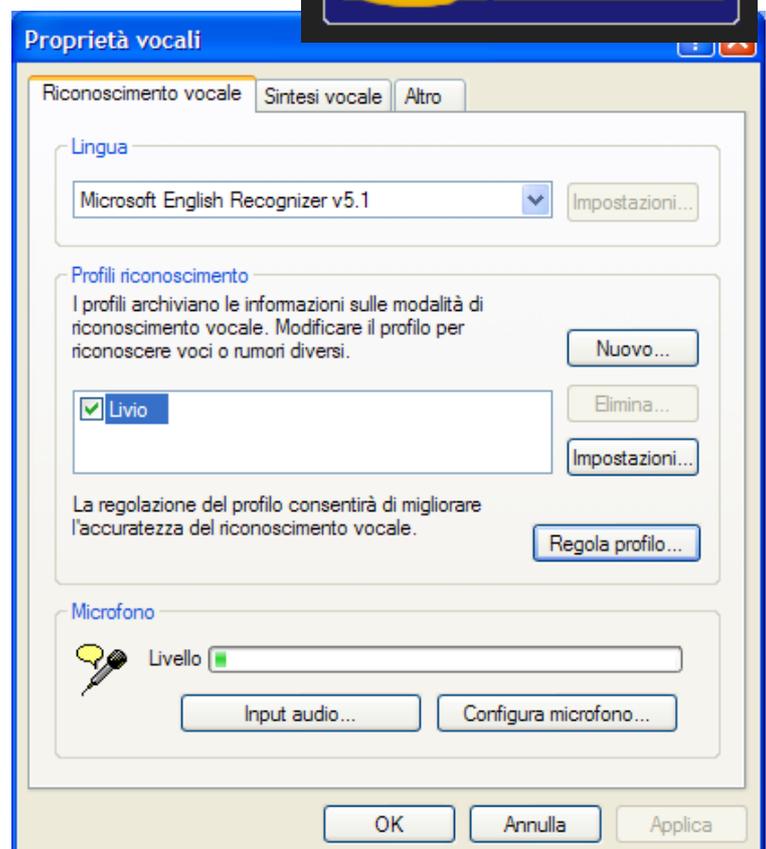
Poi premere "Nuovo", mettere un nome di utente che sia diverso da tutti quelli esistenti e premere "Fine" (**senza fare "avanti" e quindi senza proseguire con gli apprendimenti guidati**)

Ora esiste il nuovo utente ed e' già' al primo posto e con il segno di spunta verde.

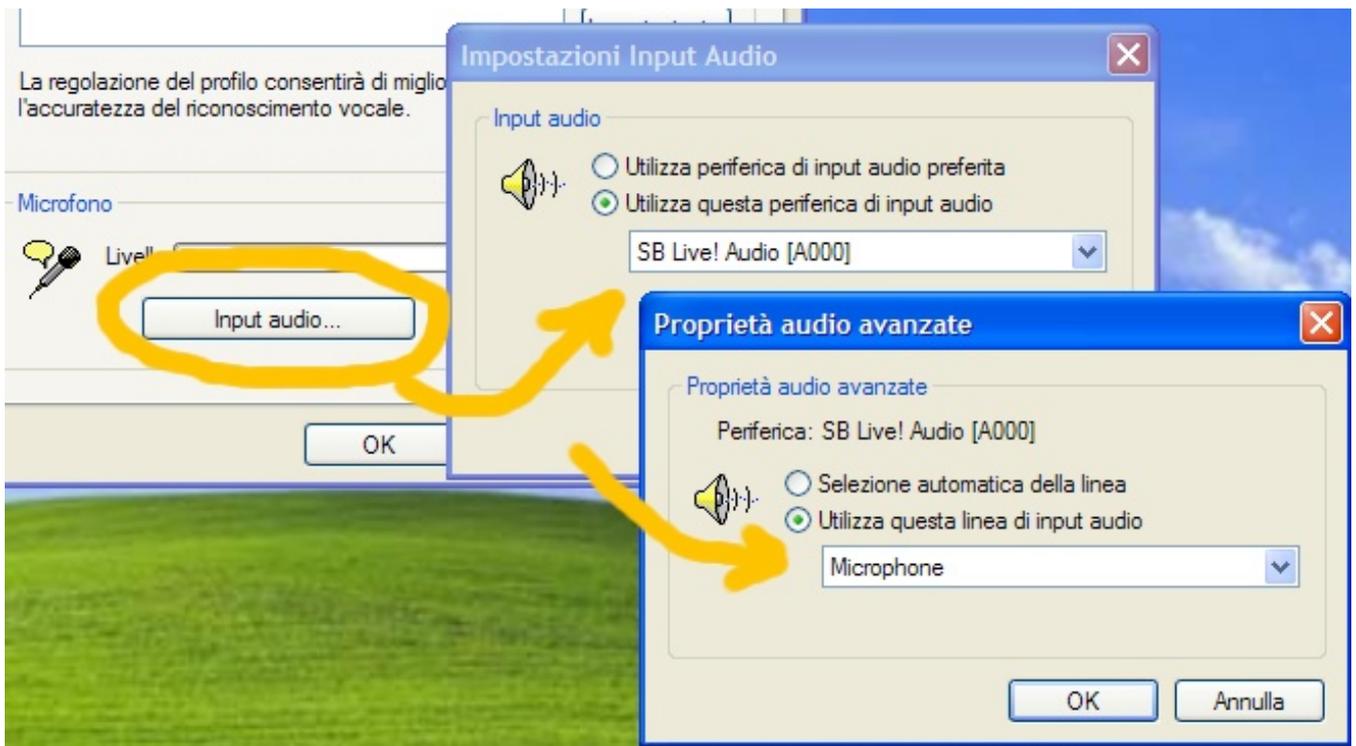
Selezionare l'altro, quello vecchio, e premere il tasto "Elimina", poi "Applica" e infine "OK"

Se vogliamo potremmo riaprire un momento questo pannello per dare una controllata che l'utente sia quello giusto.

In questo pannello **non si devono usare le procedure** "Impostazioni", "Regola profilo" e "Configura microfono" perché eseguono apprendimenti lunghissimi da eseguire, inutili e anche dannosi per la configurazione del nostro utente. In caso di errori ricominciare da capo con un utente vergine.



Prima di chiudere definitivamente questo pannello magari e' il caso di dare una occhiata a "Input audio"



Per essere sicuri che l'audio sia preso direttamente dal microfono e' una buona idea imporre "Utilizza questa periferica di input audio" e **scegliere la scheda audio e il microfono da usare** come mostrato qui sopra.

Tenere aperto il mixer audio per controllare l'apprendimento

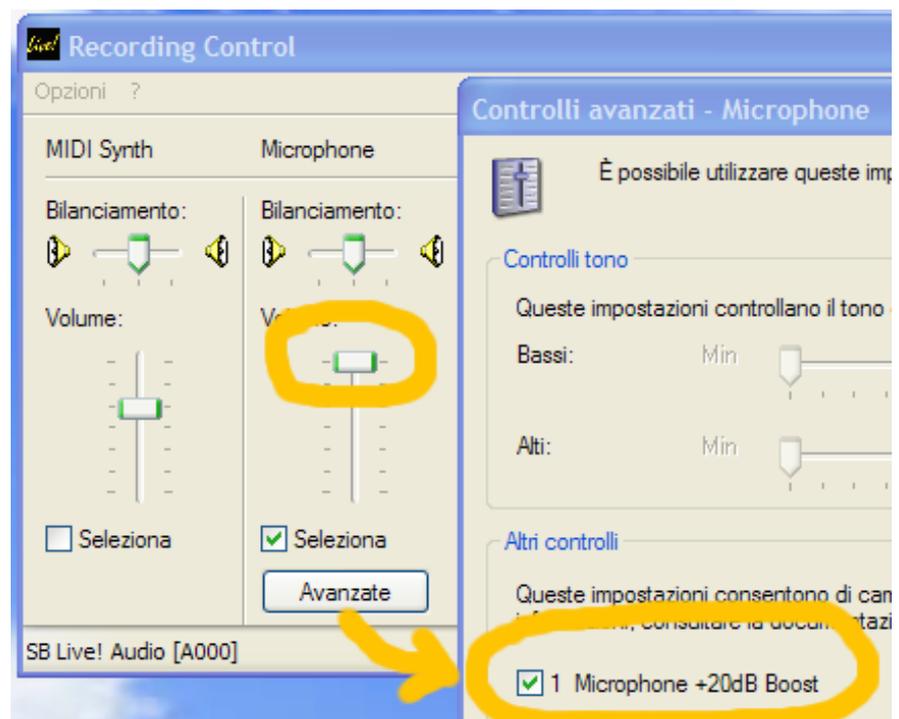
Per vedere gli effetti di quel che faremo in seguito potrebbe essere una buona idea aprire il mixer audio e tenerlo aperto. Premere il tasto "Input audio" e poi premere il tasto "Volume"

Nei casi normali il mixer audio dovrebbe essere regolato come mostrato qui di fianco. Cioè con il microfono tutto al massimo e con la preamplificazione di 20dB abilitata

Solitamente, questo non basta ancora e si deve parlare abbastanza vicino al microfono e non troppo piano, altri 10 dB di guadagno non guasterebbero...

Ma non possiamo fare queste regolazioni sul mixer direttamente perché le sue regolazioni vengono poi modificate dal riconoscimento vocale.

Per cui terremo aperto il mixer fino alla fine dell'apprendimento ma lo useremo solo per controllare che il volume resti al massimo.

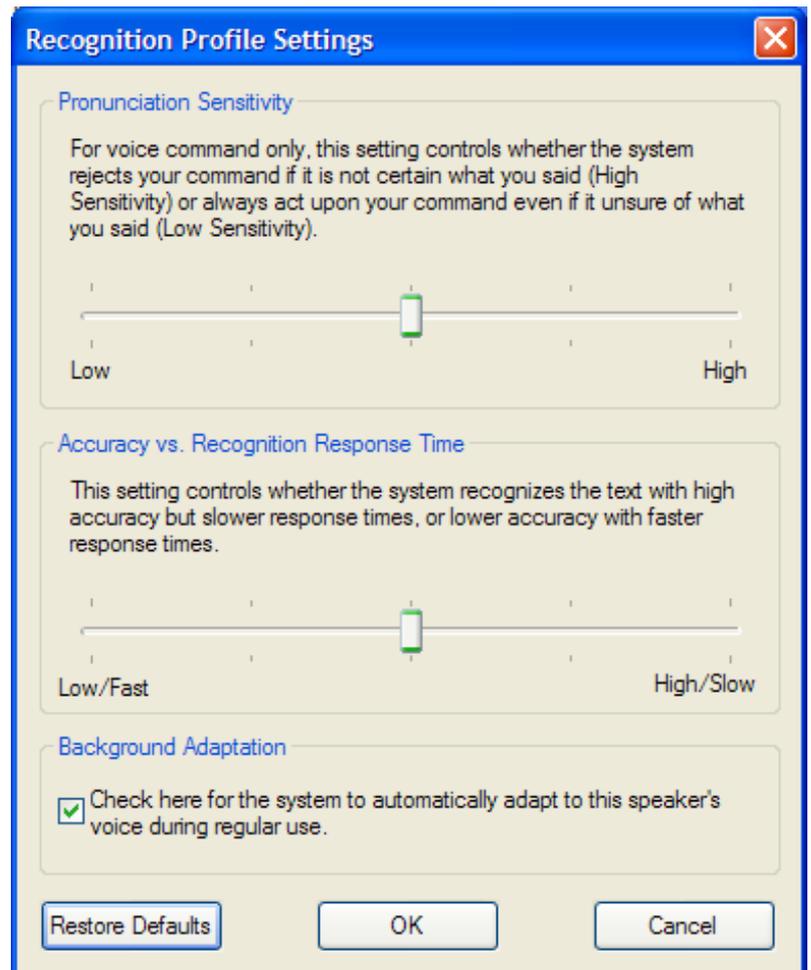


Proprietà del riconoscimento

Nel pannello "Voice input" premere il tasto "Recog. props" che apre il pannello "Recognition profile settings"



Regolare i due cursori a meta' corsa....



...togliere il segno di spunta verde da "Background adaptation"

e infine premere OK.

E' meglio non usare "Background adaptation" perché altrimenti il riconoscimento continua a cambiare nel tempo e non si e' mai sicuri di come risponde. Dopo un certo tempo alcuni comandi finiscono per prendere il sopravvento e altri diventano sordi.

Regolazione del volume di ingresso del microfono

Aprire il "Microphone wizard" premendo il tasto "Mic. Training"



Seguire le istruzioni parlando con voce normale, non troppo forte e non troppo vicino al microfono.

Durante questa regolazione controllare nel mixer che il volume del microfono resti al massimo, se capita di parlare troppo forte e il volume si abbassa si deve premere il tasto "Indietro", nel pannello del microfono e poi nuovamente il tasto "Avanti", così facendo il volume viene ripristinato al massimo.

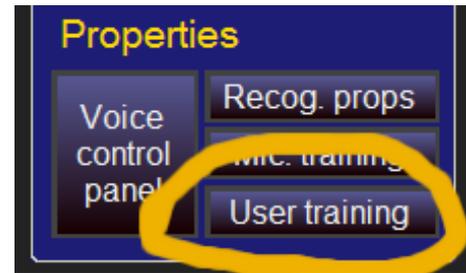
Quando si è soddisfatti chiudere il "Microphone wizard" e anche il mixer che non serve più.



Apprendimento dei comandi di FreeTW

Fortunatamente FreeTW usa solo una ventina di comandi per cui questa fase sarà velocissima.

Questa fase potrà essere anche ripetuta più volte, anche a distanza di tempo e magari parlando ogni volta in modo leggermente diverso oppure un po' più piano o più forte o a distanza diversa dal microfono. Ogni volta il sistema imparerà qualche sfumatura in più e migliorerà il riconoscimento ma attenzione a non fargli imparare anche rumori di macchine che passano o a battere sul microfono durante questa fase.



Purtroppo non è possibile azzerare l'apprendimento per cui se, per un qualunque motivo, roviniamo un utente e non riconosce più bene i comandi dovremo farne uno nuovo e ricominciare da capo.

Premere "Avanti"

Se è la prima volta che usiamo questo utente dovremo scegliere "Maschio" o "Femmina" e "Meno di 13 anni" o "Più di 13 anni"



Poi si pronunciano tutti i comandi con voce normale.

Se si fa qualche errore basta premere "Indietro" e ripronunciare il comando.

Se ci si ferma per troppo tempo avviene una pausa automatica e si deve premere "Resume"

Quando si arriva all'ultimo comando "nine" l'apprendimento viene salvato automaticamente, possiamo chiudere tutto e abbiamo finito.

Microfoni

I microfoni Shot-gun

Un microfono direttivo potrebbe migliorare moltissimo il riconoscimento vocale, l'ideale sarebbe uno "Shot-gun" di quelli che si usano per le riprese dell'audio dei film come quello mostrato qui di seguito.

Purtroppo questi microfoni devono essere molto lunghi altrimenti sono poco direttivi e non vale la pena di usarli.

Gli "shot-gun" sono ingombranti e costano dai 150 ai 300 euro.

Pero' un microfono del genere, puntato verso chi parla e tenuto a non più' di un metro di distanza potrebbe migliorare moltissimo il riconoscimento della voce.



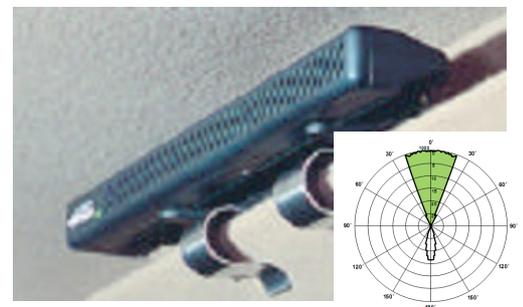
Potrebbe essere interessante il Telex M-60 visibile qui di fianco che costa abbastanza poco (circa 20 euro)

Questo modello, pur essendo piccolo, e' abbastanza direzionale da dare sensibili miglioramenti.

Gli array direzionali di microfoni

Si possono trovare, ad esempio, presso www.AndreaElectronics.com

Quelli con solo due microfoni sono poco direttivi, secondo me l'unico che potrebbe andare bene e' il DA-350 (che costa circa 50 euro)
Purtroppo il DA-350 e' fatto per essere usato in auto e quindi deve essere alimentato a 12 v con qualche scomodo adattatore...



I microfoni direzionali parabolici

Una parabola trasparente "Telinga", fatta per registrare gli uccelli, con un microfono electret nel fuoco sarebbe il massimo. Si potrebbe piazzarla in un angolo della stanza e puntarla sul divano.

Peccato che la parabola, da sola, costi 80 euro e che si trovi solo in Svezia e che i prodotti similari, reperibili su eBay, siano troppo piccoli e inefficienti.



Indirizzi di venditori di microfoni

Se si vuole dare un'occhiata a quali microfoni esistono questo sito e' uno dei migliori ma eviterei di comprare qui un microfono perché costano tutti moltissimo.

<https://www.emicrophones.com/>

Questi invece sono siti italiani con buoni prezzi :

http://www.monclick.it/vendita_cuffie-microfoni-pc

http://www.agecompc.it/products.php?id_cat=cat_ricerca&page=2&search_string=microfono

Sono da preferire i modello USB, magari un Trust da 20 euro oppure un Logitech da 24 euro in modo da non avere la pila da cambiare periodicamente.

Sono invece da scartare tutti i modelli da mettere in testa perché o hanno il filo, che poi da fastidio, oppure hanno un collegamento via radio ma allora costano moltissimo.

Microfoni non direzionali, ma preamplificati

I normali microfoni per PC (che sono degli electret, non direzionali, senza batteria e alimentati dal pc stesso) vanno bene se si parla a non più' di dieci centimetri dal microfono e con voce normale.

Per chi preferisce parlare piuttosto piano (e di conseguenza tiene anche l'audio non molto alto) un microfono preamplificato, con guadagno di dieci o quindici decibel, può fare miracoli.

Normalmente i microfoni preamplificati contengono una batteria da cambiare ogni tanto e questo potrebbe essere un fastidio.

Un microfono preamplificato senza batteria

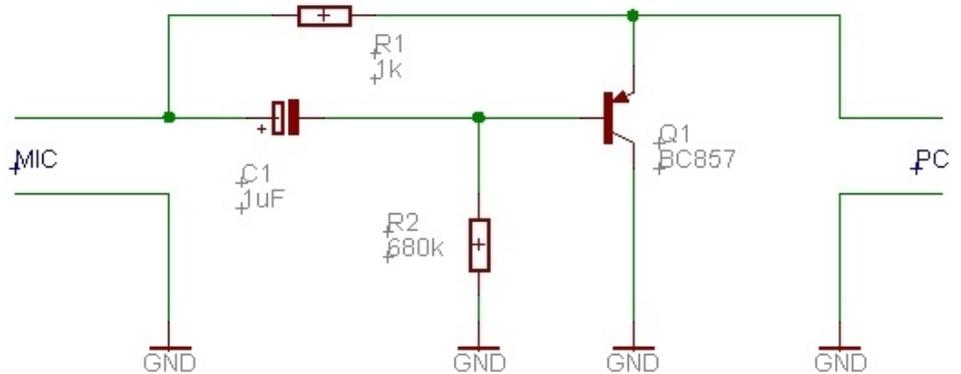
Per evitare il fastidio di dover cambiare periodicamente la batteria e anche per evitare l'ingombro e le complicazioni meccaniche dovute al portatile, esiste una soluzione molto comoda ma che non si trova in commercio. Chi fosse in grado di fare piccoli lavori di elettronica potrebbe costruirselo.

Caratteristiche :

Guadagno circa 12 dB
(minimo 8, massimo 14 a seconda della scheda audio)

Rumore di fondo bassissimo
(indistinguibile dal rumore di fondo del microfono stesso)

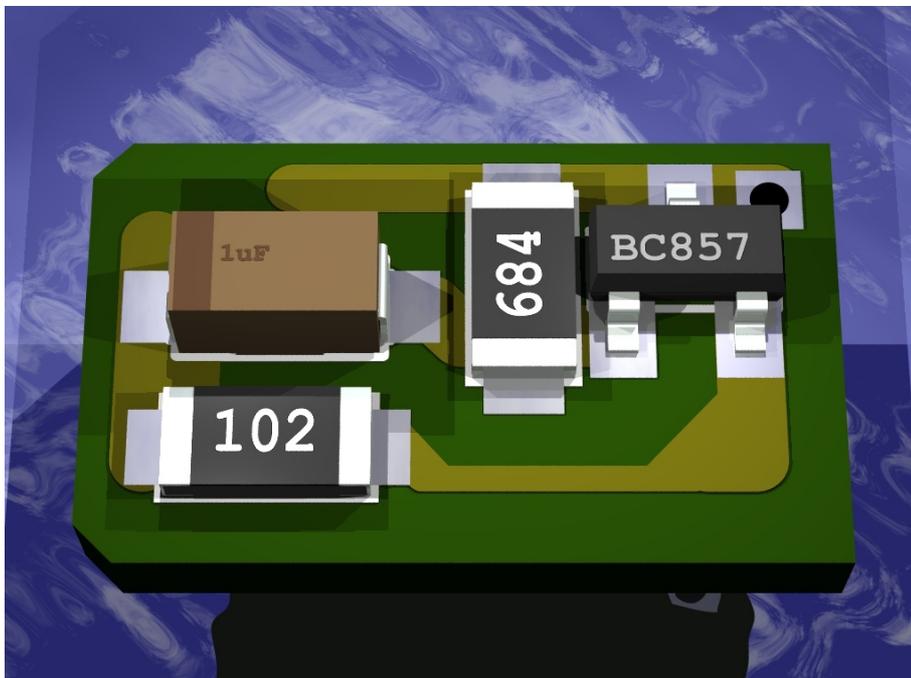
Banda passante da 10Hz a 100Khz



I componenti non sono critici, al posto del BC857 si può usare un BC307 o un qualunque altro PNP per piccoli segnali.

E' possibile fare un piccolo montaggio in aria oppure si può fare un circuito stampato con vetronite a doppia faccia. Sulla faccia inferiore, che qui non e' visibile, il rame deve essere un piano continuo.

Sulla sinistra si salda il microfono (un microfono electret da 6mm) con il negativo sulla faccia inferiore e il positivo sulla giunzione della resistenza e del condensatore.



In questo foro si inserisce un filo di piccolo diametro e lo si salda sopra e sotto.

Da qui parte il cavo schermato che va al PC. Si deve saldare la calza del cavo al rame sulla faccia inferiore e il filo centrale alla piazzola triangolare.

(dimensioni 6 x11 mm)

Se si vuole eliminare ogni possibilità di disturbi e ronzii e bene racchiudere il tutto (anche il microfono e la parte terminale spellata del cavo schermato) in un tubetto di rame da 6mm di diametro interno, lungo qualche centimetro, e saldare a questo tubetto il filo che e' saldato al buco dello stampato.

Circuito inedito - (C) Livio 2008

Usare un motore di riconoscimento diverso

Attualmente FreeTW usa "SAPI 5.1" (speech api versione 5.1 di Microsoft) per il riconoscimento vocale e anche per la sintesi vocale.

SAPI 5.1 non ha il riconoscimento in italiano e non funziona su Vista per cui si potrebbe pensare di usare qualcos'altro, anche a pagamento. Una lunga ricerca ha individuato "Nuance Dragon NaturalSpeaking 9" come il migliore e forse l'unico con le caratteristiche necessarie.

Purtroppo se si vuole usarlo e' necessario comprare la versione completa che e' l'unica in grado di usare vocabolari limitati e personalizzabili e questa versione costa moltissimo.

Risultati delle prove

Le prove hanno evidenziato che "Dragon" funziona meglio del motore di Microsoft quando si tratta di riconoscere un testo sotto dettatura e si tratta quindi di scegliere tra decine di migliaia di parole.

Quando invece le alternative sono solo una ventina, come nel nostro caso, i due motori si equivalgono e i vantaggi si riducono a uno solo, la possibilità di usare parole italiane.

In compenso gli svantaggi sono molti:

- il costo, 299 euro sono veramente troppi per fare le stesse cose che si fanno già benissimo con il telecomando

- il tempo, le complicazioni e i problemi che si tira dietro una installazione da più di un giga byte

- le difficoltà di impostazione e di uso (e' necessario caricare un dizionario personalizzato etc...)

Per chi vuole sperimentare

Comunque chi volesse fare delle prove può andare presso i seguenti indirizzi e magari prendere la versione meno costosa che va benissimo per fare esperimenti e rendersi conto di come funziona.

<http://www.nuance.com/products/>

http://www.nuancestore.com/v2.0-img/operations/scansoft/site/393027/it/dns-talk_italy.html

Valutazioni finali

Questo documento parla di così tanti problemi e regolazioni che qualcuno potrebbe farsi una idea del tutto negativa sui comandi vocali.

In realtà se si usa un microfono ben funzionante, si regola tutto bene, si parla vicino al microfono e abbastanza chiaramente il riconoscimento è assicurato **al cento per cento** e si può stare certi che **non si sbaglia mai un comando**.

Naturalmente col tempo ci si abitua e più il sistema riconosce bene e più si finisce per parlare piano e male... alla fine ci si stabilizza a parlare abbastanza piano e non dover ripetere "*quasi mai*" i comandi.