



# APLICAÇÕES MILITARES DA TECNOLOGIA DA VOZ(\*)

Sidney Cerqueira Bispo dos Santos

---

Mostrar a importância da pesquisa de voz, e algumas de suas aplicações na área militar, é o objetivo a que se propõe este artigo.

---

## INTRODUÇÃO

**N**unca um conflito armado foi tão abundantemente noticiado quanto a guerra ocorrida recentemente no Golfo Pérsico. Entretanto as informações que chegaram ao grande público foram fortemente censuradas. Devido às características próprias do antagonismo, que envolvia nações com comunicações altamente sofisticadas de um lado e exatamente o oposto do outro, vários aspectos não receberam a devida atenção por parte da imprensa especializada. Um desses aspectos, fundamental em um estado de beligerância, é justamente a segurança das comunicações. A rapidez

com que se trocam informações, a imunidade ao ruído existente no teatro de operações e à interferência eletromagnética intencional, o sigilo das mensagens e a escuta inimiga são considerações que estão sempre presentes. As pesquisas nessas áreas são enormes e ininterruptas, pois delas depende, fundamentalmente, o sucesso de qualquer operação militar.

Nos países do primeiro mundo, os investimentos nas áreas referentes à segurança são bastantes elevados. Dentre essas áreas uma se destaca, pelas elevadas cifras envolvidas (bilhões de dólares): o processamento digital dos sinais de voz. A importância militar dessas pesquisas aumenta, quando se verifica que os sistemas desenvolvidos baseados na parametrização da voz

(\*) Selecionado pelo PADECEME

consideram, além das particularidades lingüísticas, fatores comportamentais, regionalismos e fatores fisiológicos.

Isto quer dizer que um sistema desenvolvido baseado em parâmetros da língua inglesa, quando utilizado, por exemplo, por um brasileiro, não funcionará satisfatoriamente, mesmo que este fale fluentemente o inglês, pois as diferenças fisiológicas irão influir.

Existe uma relação biunívoca entre o aparelho fonador e a língua. Uma criança ao balbuciar as primeiras sílabas, tentando imitar um adulto, condiciona seu aparelho fonador, ainda em formação, a se estruturar de tal forma que os fonemas daquele idioma ao serem pronunciados resultem num menor esforço para os órgãos envolvidos na fala. Dessa maneira, torna-se difícil para um adulto, cujo aparelho fonador está totalmente desenvolvido, falar uma outra língua que possua fonemas diferentes, pois, além da necessidade de aprender a posicionar os articuladores (língua, queixo, bochechas etc.) para a formação de novos sons, haverá o desgaste físico — os músculos serão forçados a produzir fonemas para os quais não foram estruturados.

O povo brasileiro possui idiossincrasias bastante características. O português falado no Brasil possui nuances que o diferenciam, quando comparado ao português falado em Portugal ou outros países que falam a língua portuguesa. E, mesmo esta, possui fonemas que são únicos. Portanto, é dentro desse contexto que deve ser encarada a pesquisa de voz no Brasil.

Independente dos avanços tecnoló-

gicos alcançados pelo primeiro mundo, a pesquisa dos sinais de voz deve ser incentivada nos grandes centros de estudo, porque se trata de desenvolver sistemas baseados em nossa língua e, principalmente, em órgãos militares, devido à sua grande utilidade em sistemas de segurança.

## APLICAÇÕES MILITARES DA TECNOLOGIA DA VOZ

As aplicações militares dos sinais de voz podem ser distribuídas em cinco grandes áreas: reconhecimento da fala; reconhecimento do locutor; síntese de voz; quantização vetorial; e processamento de voz distorcida.

Em todas essas aplicações, o sinal deverá sofrer alguma forma de processamento. Tipicamente, se extraem determinados parâmetros que permitem a caracterização do sinal de entrada utilizando uma quantidade menor de dados. Esses parâmetros dependerão da configuração do sistema e do tipo da aplicação.

### Reconhecimento da Fala

No *reconhecimento da fala*, o que se deseja é determinar, a partir do sinal de voz digitalizado, o que foi dito pelo locutor.

O principal objetivo militar do *reconhecimento da fala* é permitir que um piloto, ou um operador de equipamentos ou máquinas, realize tarefas sem distrações manuais ou visuais.



consideram, além das particularidades lingüísticas, fatores comportamentais, regionalismos e fatores fisiológicos.

Isto quer dizer que um sistema desenvolvido baseado em parâmetros da língua inglesa, quando utilizado, por exemplo, por um brasileiro, não funcionará satisfatoriamente, mesmo que este fale fluentemente o inglês, pois as diferenças fisiológicas irão influir.

Existe uma relação biunívoca entre o aparelho fonador e a língua. Uma criança ao balbuciar as primeiras sílabas, tentando imitar um adulto, condiciona seu aparelho fonador, ainda em formação, a se estruturar de tal forma que os fonemas daquele idioma ao serem pronunciados resultem num menor esforço para os órgãos envolvidos na fala. Dessa maneira, torna-se difícil para um adulto, cujo aparelho fonador está totalmente desenvolvido, falar uma outra língua que possua fonemas diferentes, pois, além da necessidade de aprender a posicionar os articuladores (língua, queixo, bochechas etc.) para a formação de novos sons, haverá o desgaste físico — os músculos serão forçados a produzir fonemas para os quais não foram estruturados.

O povo brasileiro possui idiossincrasias bastante características. O português falado no Brasil possui nuances que o diferenciam, quando comparado ao português falado em Portugal ou outros países que falam a língua portuguesa. E, mesmo esta, possui fonemas que são únicos. Portanto, é dentro desse contexto que deve ser encarada a pesquisa de voz no Brasil.

Independente dos avanços tecnoló-

gicos alcançados pelo primeiro mundo, a pesquisa dos sinais de voz deve ser incentivada nos grandes centros de estudo, porque se trata de desenvolver sistemas baseados em nossa língua e, principalmente, em órgãos militares, devido à sua grande utilidade em sistemas de segurança.

## APLICAÇÕES MILITARES DA TECNOLOGIA DA VOZ

As aplicações militares dos sinais de voz podem ser distribuídas em cinco grandes áreas: reconhecimento da fala; reconhecimento do locutor; síntese de voz; quantização vetorial; e processamento de voz distorcida.

Em todas essas aplicações, o sinal deverá sofrer alguma forma de processamento. Tipicamente, se extraem determinados parâmetros que permitem a caracterização do sinal de entrada utilizando uma quantidade menor de dados. Esses parâmetros dependerão da configuração do sistema e do tipo da aplicação.

### Reconhecimento da Fala

No *reconhecimento da fala*, o que se deseja é determinar, a partir do sinal de voz digitalizado, o que foi dito pelo locutor.

O principal objetivo militar do *reconhecimento da fala* é permitir que um piloto, ou um operador de equipamentos ou máquinas, realize tarefas sem distrações manuais ou visuais.



Dentre as principais aplicações podemos citar:

- ativação e controle de sistemas de armas;
- operação de sistemas controlados remotamente tais como robôs ou submersíveis;
- controle de várias funções em tanques, aviões, submarinos, navios etc.

Existem ainda uma grande vantagem em se utilizar um sistema de *reconhecimento da fala* como entrada de dados em um computador. Como o teclado é o principal instrumento utilizado atualmente, cada terminal exige, além do teclado e vídeo, microprocessadores para controle do tráfego de informações, enquanto que, em um sistema que utilize a tecnologia da voz, cada telefone se transformaria em um terminal com todas as facilidades e vantagens decorrentes, além da rapidez que se conseguiria, pelo fato de que os operadores não necessitariam de treinamento especializado.

### Reconhecimento do Locutor

Enquanto no *reconhecimento da fala* o que interessa é o que está sendo dito, no *reconhecimento do locutor* o que interessa é quem está dizendo algo.

Pode-se distinguir dois tipos de tarefas que utilizam técnicas diferentes sob o título de *reconhecimento do locutor*: *verificação de locutor* e *identificação de locutor*.

Um aspecto interessante na distinção entre essas duas tarefas é que, enquanto na *verificação de locutor*, o locutor é cooperativo, ou seja, ele quer

ser reconhecido, na *identificação* pode não ser, ou seja, o locutor normalmente tentará disfarçar a voz.

Sistemas de *verificação de locutor* podem ser aplicados principalmente em sistemas de segurança de qualquer setor militar restringindo o acesso a áreas controladas. Um sistema desse tipo pode complementar, ou mesmo eliminar, o posto de guarda nas entradas de áreas restritas. Também pode ser utilizado como segurança inicial em um sistema de controle de armas acionado por voz.

Sistemas de *identificação de locutor* podem ser aplicados mais apropriadamente na área da inteligência militar, onde seja importante determinar a identificação de uma pessoa, a partir somente de sua voz, independente do que esteja sendo falado.

### Síntese de Voz

A *síntese de voz*, além da sua óbvia utilização em sistemas de resposta, possui uma importância militar muito grande, pois pode permitir um maior grau de liberdade na utilização de equipamentos e veículos.

Normalmente, pilotos e operadores de sistemas de armas ou equipamentos necessitam ter, quando em combate, toda sua atenção voltada para o inimigo. Dessa forma, os inúmeros dispositivos, mostradores e painéis, que informam o funcionamento ou, mesmo, uma correção na direção ou na operação do equipamento, se tornam, no mínimo, uma preocupação a mais. Assim, um sistema de alerta utilizando voz sintetizada, dispensaria o piloto ou



operador de algumas inspeções visuais, aumentando bastante sua eficiência.

## Quantização Vetorial

A situação hoje em dia é tal, que o único sistema realmente seguro (indecifrável) é o baseado no uso de códigos pseudo-aleatórios. Dentro desse contexto, a *quantização vetorial* é de fundamental importância na segurança das comunicações, pois, além da codificação natural do processo (o que torna impossível a sua decodificação por quem não conhece os parâmetros), permite a redução do tempo de transmissão (pode-se transmitir por salvas), aumenta a velocidade de transmissão e reduz a banda ocupada pelo sinal (fundamental para utilização de satélites).

Esse processo ainda proporciona grande imunidade ao ruído interferente, intencional ou não, pois permite a introdução de redundâncias inteligentes.

## Processamento de Voz Distorcida

Um teatro de operações é extremamente hostil aos sistemas de comunicações por voz, na forma de ruído e distorções.

As fontes de ruído podem ser várias. Por exemplo, deixando de lado o ruído ocasionado por interferência intencional, temos os ruídos naturais do meio nos sistemas instalados em aviões, helicópteros, tanques etc. Para

complicar ainda mais, quando um indivíduo permanece exposto a um nível alto de ruído durante um certo tempo, ele tende a aumentar, tanto o volume, quanto a frequência fundamental da sua voz, modificando o sinal.

As *distorções no sistema* podem ser introduzidas em diversos níveis. Por exemplo, o primeiro componente no processamento do sinal de voz é o microfone. Nele as distorções podem ter diversas causas, tais como resposta de frequência inadequada, descasamento na configuração capacete/microfone ou, simplesmente, mau funcionamento. Outras formas de *distorções* podem ocorrer se o locutor estiver utilizando máscaras (de piloto, contra gases, de mergulhador etc.). A máscara pode aumentar a ressonância do trato vocal, dificultar a respiração, ou mesmo restringir o movimento normal dos articuladores.

*Distorções na voz* também podem ser causadas por meios hipobáricos (alta altitude e baixa pressão) e hiperbáricos (grande profundidade e alta pressão).

Além dos problemas causados pelo ruído e distorções, existe ainda um outro fator que pode contribuir para a degradação do sinal de voz: a situação física e psicológica do locutor. Se este está em situação de *stress* (sob força G, temperaturas extremas, ferido ou fatigado) a voz pode ser afetada. O *stress* emocional ou psicológico pode ainda ser causado por situações de emergência, percepção de perigo, sobrecarga mental ou medo.

Assim, é importante que os efeitos desses problemas na voz sejam consi-



derados no projeto de sistemas para aplicações militares.

### A PESQUISA DA TECNOLOGIA DA VOZ NO INSTITUTO MILITAR DE ENGENHARIA

A Seção de Engenharia Elétrica do IME possui, na área de concentração *processamento de sinais*, três linhas de pesquisas (LP): processamento da voz; processamento de imagens; e processamento de sinais aplicado às comunicações e sistemas de acompanhamento.

A LP *processamento da voz* foi ativada no IME pela primeira vez em 1977. Desde então, volta-se principalmente às atividades de *reconhecimento de locutores*, *reconhecimento da fala*, *quantização vetorial* e *síntese da voz*. Frutos dessas pesquisas, já foram geradas 11 teses, 9 projetos de fim de curso e 5 trabalhos de iniciação à pesquisa.

Atualmente, mantendo a tradição em pesquisa de ponta, estuda a utilização de modelos de redes neurais artificiais (redes que simulam o

funcionamento do cérebro humano) nos sistemas citados anteriormente.

### REFERÊNCIAS BIBLIOGRÁFICAS

- SANTOS, Sidney Cerqueira Bispo dos, *Fonemas Nasalados na Verificação Automática de Locutores*, 9º SBT, USP, São Paulo - SP, setembro de 1991.
- MOORE, C.A., MOORE, R.D., *Pilot-Aircraft Interface — The Voice Channel*, Proceedings Of International Speech Tech '87, maio 1987. (Abstract).
- SMITH, K., BOWEN, W.J.C., *The Development Of Speech Recognition For Military Aviation*, Proceedings of International Speech Tech '87, maio 1987. (Abstract).
- WILSON, J., *Military Speech Recognisers*, Proceedings Of International Speech Tech '87, maio 1987. (Abstract).
- HAAKE, J., BENSON, P., KOBLE, H., *Automatic Speech Understanding For Naval Battle Management*, Proceedings Of International Speech Tech '87, maio 1987. (Abstract).
- MAYER, M.S.M FUNG, T.Y., *Communications In The High Noise Enviroments Of Tracked Combat Vehicles*, Proceedings Of International Speech Tech '87, maio 1987. (Abstract).
- JAMES W. Hicks, Jr., *Military Applications For Voice Technology*, Electronic Show and Convention, Los Angeles — CA, 1984.



*SIDNEY CERQUEIRA BISPO DOS SANTOS é major do Quadro de Engenheiros Militares do Exército. Concluiu a AMAN (Comunicações) em 1977. Diplomado em Engenharia de Comunicações pelo IME em 1985, obteve o grau de Mestre em Ciências — Engenharia Elétrica em 1989 pelo mesmo Instituto. Atualmente é Chefe do Laboratório de Processamento de Sinais e professor/pesquisador da Seção de Eletricidade e Eletrônica do IME, trabalhando na área de comunicação homem-máquina por voz.*